

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:12 ; Search time 38.7752 Seconds

(Without alignments)
711.353 Million cell updates/sec

Title: US-09-917-791-21

Perfect score: 1071

Sequence: 1 IKVNMNLFSPSEDFNTND.....NIGNMLYKDDFVGLIFSGA 207

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1071	100.0	207	22	AAG79295	Amino acid sequence
2	1071	100.0	407	21	AAV77141	Native botulinum n
3	1071	100.0	413	23	ABG69074	Botulinum toxin hea
4	1071	100.0	413	23	ABG69074	Botulinum neurotox
5	1071	100.0	423	21	AAAB36302	C. botulinum BONT/
6	1071	100.0	837	21	AAV77140	Native botulinum n
7	1071	100.0	861	23	ABG69075	Botulinum neurotox
8	1071	100.0	871	19	AAW56019	Recombinant botuli
9	1071	100.0	871	19	AAW56007	Recombinant botuli
10	1071	100.0	871	19	AAW56008	Botulinum neurotox

11	1071	100.0	873	19	AAW56016	Recombinant botuli
12	1071	100.0	875	19	AAW56009	Recombinant botuli
13	1071	100.0	878	19	AAW56010	Recombinant botuli
14	1071	100.0	894	19	AAW56015	Recombinant botuli
15	1071	100.0	907	19	AAW56012	Recombinant botuli
16	1071	100.0	953	19	AAW56011	Recombinant botuli
17	1071	100.0	1013	19	AAW56013	Recombinant botuli
18	1071	100.0	1067	21	AAV93307	A manganese supero
19	1071	100.0	1092	21	AAV93310	A manganese supero
20	1071	100.0	1295	23	AAU99339	Clostridium botuli
21	1071	100.0	1296	17	AAW5010	C. botulinum type
22	1054.5	98.5	847	22	AAW5010	Botulinum toxin hea
23	377	35.2	410	22	AAW5010	Botulinum toxin hea
24	368.5	34.4	804	23	ABG69083	Botulinum toxin hea
25	368.5	34.4	804	23	ABG69083	Botulinum toxin hea
26	368.5	34.2	413	22	AAW5017	Botulinum toxin hea
27	368.5	34.2	848	22	AAW5017	Botulinum toxin hea
28	368.5	34.2	852	23	AAW5017	Botulinum toxin hea
29	368.5	34.2	858	19	AAW5018	Recombinant botuli
30	368.5	34.2	1070	21	AAV93308	A manganese supero
31	368.5	34.2	1095	21	AAV93311	A manganese supero
32	368.5	34.2	1169	19	AAW5017	Recombinant botuli
33	368.5	34.0	848	23	ABG69087	Botulinum neuroto
34	364.5	34.0	1059	21	AAV93309	A manganese supero
35	364.5	34.0	1084	21	AAV93312	A manganese supero
36	364.5	34.0	1291	19	AAW68392	Clostridium botuli
37	316	29.5	408	22	AAW5011	Botulinum toxin hea
38	316	29.5	858	23	ABG69085	Botulinum neurotox
39	290.5	27.1	1315	22	AAW5011	Clostridium tetrat
40	290.5	27.1	386	22	AAW5011	Botulinum toxin hea
41	290.5	27.1	824	23	AAW5011	Botulinum neurotox
42	263	24.6	399	22	AAW5011	Botulinum toxin hea
43	263	24.6	811	23	ABG69079	Botulinum neurotox
44	263	24.6	1291	20	AAW5814	Non-toxic modified
45	116.5	10.9	548	19	AAW56014	Recombinant botuli

ALIGNMENTS

RESULT 1
AAG79295
AAG79295 standard; Protein: 207 AA.
AC AAG79295;
DT 03-JAN-2002 (first entry)
XX Amino acid sequence of botulinum neurotoxin (BONT) fragment.
XX Neurotoxin; BONT; vaccine; botulism.
XX Clostridium botulinum.
XX US6287566-B1.
XX 11-SEP-2001.
XX 19-MAY-1995; 95US-0446114.
XX 19-MAY-1995; 95US-0446114.
XX (USSA) US SEC OF ARMY.
XX Dertzbauhg MT;
XX WPI; 2001-615462/71.
XX New protective epitopes of neurotoxin of Clostridium botulinum, useful
XX for developing vaccines against neurotoxins of Clostridium botulinum
XX and for identifying protective antibodies
XX Claim 1; Column 25; 14pp; English.

applicant

XX The present sequence represents a fragment of the Clostridium botulinum
 CC neurotoxin (BoNT). It was produced by amplifying overlapping fragments
 CC of the BoNT gene. The amplified fragments were cloned expressed to
 CC identify immunogenic polypeptides which are capable of giving rise to
 CC protective antibodies. The BoNT polypeptide fragment are useful as
 CC vaccines, for immunizing against botulism, and as diagnostic agents
 CC to identify protective antibodies.

XX
 SQ Sequence 207 AA:

Query Match 100.0%; Score 1071; DB 22; Length 207;
 Best Local Similarity 100.0%; Pred. No. 1.2e-99;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMWDLFFSPSEDNFTNDLNGEITSDTNIIEAEENISDLIOQYLYTFNFNEPEN 60
 DB 1 IKVNMWDLFFSPSEDNFTNDLNGEITSDTNIIEAEENISDLIOQYLYTFNFNEPEN 60
 QY 61 ISENLSSDIIGOLELMPNIEERPNGKKEYELDKYTFMFLRAQEFHGKSRALNNSVNE 120
 DB 61 ISENLSSDIIGOLELMPNIEERPNGKKEYELDKYTFMFLRAQEFHGKSRALNNSVNE 120
 QY 121 ALLNPSRVYTFPSSDYKKVKNKATEAMFLGWVEQLVYDFTDETSEVSTTDKIADITITII 180
 DB 121 ALLNPSRVYTFPSSDYKKVKNKATEAMFLGWVEQLVYDFTDETSEVSTTDKIADITITII 180
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
 DB 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207

RESULT 2
 AAY77141
 ID AAY77141 standard; Protein; 407 AA.

XX
 AC AAY77141;

XX
 DT 08-MAY-2000 (first entry)

XX
 DE Native botulinum neurotoxin serotype A (BoNTA) N-terminal fragment (Hn).

XX
 KM Botulinum neurotoxin; heavy chain; BoNT; serotype A;
 KM N-terminal fragment; Hn; Venezuelan equine encephalitis virus replicon;
 KM VEE; botulism; vaccine; diagnosis; drug screening.

XX
 OS Clostridium botulinum.

XX
 PN WO200002524-A2.

XX
 PD 20-JAN-2000.

XX
 PF 09-JUL-1999; 99WO-US15570.

XX
 PR 10-JUL-1998; 98US-0092416.
 PR 12-MAY-1999; 99US-0133870.

XX
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.

XX
 PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
 XX WPI; 2000-160827/14.
 DR N-PSDB; AA287219.

XX
 PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
 PT toxin serotypes A-G, is used for inducing an immune response against
 PT botulinum -

XX
 PS Example 3; Page 51; 54pp; English.

XX
 CC The invention relates to novel vaccines that induce a protective immune
 CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
 CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant

CC DNA construct comprising a vector, and at least one nucleic acid
 CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
 CC serotypes A-G. In preferred embodiments of the invention, the vector is a
 CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
 CC vector results in the production of large amounts of a protein encoded by
 CC a sequence cloned into the replicon. The constructs are used to produce
 CC vaccines against botulism. The proteins can also be used as diagnostic
 CC tools for the diagnosis of botulism. The transformed host cells can be
 CC used to analyse the effectiveness of drugs and agents which inhibit toxin
 CC effects. The vaccine currently used against botulism is dangerous
 CC and expensive to produce, and contains formalin, which is very painful
 CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
 CC the 7 serotypes are represented in the formulation. The novel vaccine
 CC of overcomes these problems, as it is easily purified, and available in
 CC large quantities. It is also expressed in the lymph nodes for a better
 CC immune response. The present sequence represents the native BoNTA heavy
 CC chain N-terminal fragment (Hn) used in an exemplification of the present
 CC invention.

XX
 SQ Sequence 407 AA:

Query Match 100.0%; Score 1071; DB 21; Length 407;
 Best Local Similarity 100.0%; Pred. No. 3.2e-99;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMWDLFFSPSEDNFTNDLNGEITSDTNIIEAEENISDLIOQYLYTFNFNEPEN 60
 DB 2 IKVNMWDLFFSPSEDNFTNDLNGEITSDTNIIEAEENISDLIOQYLYTFNFNEPEN 61
 QY 61 ISENLSSDIIGOLELMPNIEERPNGKKEYELDKYTFMFLRAQEFHGKSRALNNSVNE 120
 DB 62 ISENLSSDIIGOLELMPNIEERPNGKKEYELDKYTFMFLRAQEFHGKSRALNNSVNE 121
 QY 121 ALLNPSRVYTFPSSDYKKVKNKATEAMFLGWVEQLVYDFTDETSEVSTTDKIADITITII 180
 DB 122 ALLNPSRVYTFPSSDYKKVKNKATEAMFLGWVEQLVYDFTDETSEVSTTDKIADITITII 181
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
 DB 182 PYIGPALNIGNMLYKDDFVGALIFSGA 208

RESULT 3
 AAB04168
 ID AAB04168 standard; Protein; 413 AA.

XX
 AC AAB04168;

XX
 DT 11-APR-2001 (first entry)

XX
 DE Botulinum toxin heavy chain N-terminal sequence (serotype A).

XX
 KM Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
 KM recombinant vector; antigen; immune response; vaccine; bacterium;
 KM infection.

XX
 OS Synthetic.

XX
 OS Clostridium botulinum.

XX
 PN WO200067700-A2.

XX
 PD 16-NOV-2000.

XX
 PF 12-MAY-2000; 2000WO-US12890.

XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.

PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H.
 XX
 DR WPI; 2001-016048/02.
 DR N-PSDB; AAA54590.
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX
 PS Disclosure; Fig 11b; 73pp; English.
 XX
 CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 SQ Sequence 413 AA;
 Query Match 100.0%; Score 1071; DB 22; Length 413;
 Best Local Similarity 100.0%; Pred. No. 3.3e-99;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKVNNMDLFFSPSDNFTNDLNGKEITSDTNIAEENISLDLIQOYLLFNFNDNEPEN 60
 DB 8 IKVNNMDLFFSPSDNFTNDLNGKEITSDTNIAEENISLDLIQOYLLFNFNDNEPEN 67
 QY 61 ISENLSDDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGKSRIALTNSVNE 120
 DB 68 ISENLSDDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGKSRIALTNSVNE 127
 QY 121 ALLNPSRYVTFSSDYKVKVKAATEAMFLGVOLVDFDTESEVSTTDKIADITITII 180
 DB 128 ALLNPSRYVTFSSDYKVKVKAATEAMFLGVOLVDFDTESEVSTTDKIADITITII 187
 QY 181 PYIGPALNIGMLYKDDPVGALIFSGA 207
 DB 188 PYIGPALNIGMLYKDDPVGALIFSGA 214
 RESULT 4
 AAB69074
 ID AAB69074 standard; Protein: 413 AA.
 AC AAB69074;
 XX
 XX 07-OCF-2002 (first entry)
 DE Botulinum neurotoxin light chain polypeptide #8.
 XX
 KW Botulinum neurotoxin light chain; BoNT LC; botulism; dystonia; pain;
 KW spasticity; ocular motility; facial dyskinesia; stiff-person syndrome;
 KW bladder dysfunction; segmental myoclonus; hyperkinetic disorder;
 KW cosmetic treatment; facial wrinkle; cerebral palsy; analgesic; relaxant;
 KW lower motor neuron hyperactivity; autonomic nerve function; muscular;
 KW immunostimulant; antibacterial.
 XX
 OS Clostridium botulinum.
 XX
 PN WO200236758-A2.

XX
 PD 10-MAY-2002.
 XX
 PE 06-NOV-2001; 2001WO-US47230.
 XX
 PR 06-NOV-2000; 2000US-246774P.
 PR 20-JUL-2001; 2001US-091018P.
 PR 09-AUG-2001; 2001US-311966P.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Jensen M;
 XX
 XX WPI; 2002-575192/61.
 DR N-PSDB; ABR98544.
 XX
 PT Novel nucleic acid molecule encoding botulinum neurotoxin light chain
 PT serotype A, useful for producing the neurotoxin for vaccination against
 PT botulism, comprises sequence expressible in host other than Clostridium
 PT
 XX
 PS Disclosure; Page 130-131; 166pp; English.
 XX
 CC The invention relates to a nucleic acid molecule encoding a botulinum
 CC neurotoxin light chain (BoNT LC) serotype A, where the DNA has a sequence
 CC that is expressible in a host organism other than Clostridium, or has a
 CC total A+T content that is less than about 70%. The BoNT LC protein is
 CC useful in vaccination against botulism, for eliciting protective immunity
 CC in a mammal, for treating dystonias, spasticity, pain, ocular motility,
 CC facial dyskinesias, stiff-person syndrome, bladder dysfunction, segmental
 CC myoclonus, hyperkinetic disorders, cosmetic treatment of facial wrinkles,
 CC conditions characterised by hyperactivity of the lower motor neuron, and
 CC to control autonomic nerve function or lipoe-walking due to stiff
 CC muscles common in children with cerebral palsy. The sequences are also
 CC useful for screening for botulinum neurotoxin inhibitors. This sequence
 CC represents a botulinum neurotoxin light chain serotype A protein.
 XX
 SQ Sequence 413 AA;
 Query Match 100.0%; Score 1071; DB 23; Length 413;
 Best Local Similarity 100.0%; Pred. No. 3.3e-99;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IKVNNMDLFFSPSDNFTNDLNGKEITSDTNIAEENISLDLIQOYLLFNFNDNEPEN 60
 DB 8 IKVNNMDLFFSPSDNFTNDLNGKEITSDTNIAEENISLDLIQOYLLFNFNDNEPEN 67
 QY 61 ISENLSDDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGKSRIALTNSVNE 120
 DB 68 ISENLSDDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGKSRIALTNSVNE 127
 QY 121 ALLNPSRYVTFSSDYKVKVKAATEAMFLGVOLVDFDTESEVSTTDKIADITITII 180
 DB 128 ALLNPSRYVTFSSDYKVKVKAATEAMFLGVOLVDFDTESEVSTTDKIADITITII 187
 QY 181 PYIGPALNIGMLYKDDPVGALIFSGA 207
 DB 188 PYIGPALNIGMLYKDDPVGALIFSGA 214
 RESULT 5
 AAB36302
 ID AAB36302 standard; Protein: 423 AA.
 XX
 XX AAB36302;
 AC AAB36302;
 XX
 XX 15-FEB-2001 (first entry)
 DE C. botulinum BoNT/A neurotoxin heavy chain prototoxin SEQ ID NO:8.
 XX
 KW Human; procholecystokinin; CCK A receptor; CCK B receptor;
 KW pancreatitis; antiinflammatory.
 XX

OS	Clostridium botulinum.
XX	WO20061192-A2.
PN	19-OCT-2000.
PD	
XX	
PF	06-APR-2000; 2000MO-US09142.
PP	
PR	08-APR-1999; 99US-0288326.
XX	(ALLR) ALLERGAN SALES INC.
PA	
PI	Steward LE, Sachs G, Aoki KR:
XX	WPI: 2000-679416/66.
DR	
PT	New composition for treating acute pancreatitis, comprises a pancreatic
PT	cell surface marker binding element, a translocation element that
PT	transfers polypeptide across vesicular membrane, and a therapeutic
PT	element -
XX	
PS	Disclosure; Page 28; 50pp; English.
XX	
CC	The present invention describes a composition (I) for treating acute
CC	pancreatitis. (I) comprises a first element containing a binding element
CC	that binds to a pancreatic cell surface marker, a second element
CC	containing a translocation element that facilitates polypeptide transfer
CC	across the vesicular membrane, and a third element containing a
CC	therapeutic element that inhibits enzyme secretion in pancreatic cell
CC	cryoplasm. Also described is a method for making a therapeutic
CC	polypeptide having a binding element selective for cholecystokinin (CCK)
CC	receptor by expressing within a host cell a recombinant chimeric
CC	polypeptide comprising an extein containing a therapeutic element and a
CC	translocational element, and an intein located to the carboxy terminal
CC	of extein having a cysteine, serine or threonine at its amino terminus,
CC	and contacting the extein with a synthetic peptide comprising a CCK
CC	amino acid sequence containing an amidated phenylalanine at a natural
CC	C-terminus, and a cysteine, serine or threonine at its N-terminus, and
CC	a nucleophilic reagent able to cause cleavage of the intein to form a
CC	peptide bond between the extein C-terminus and synthetic peptide
CC	N-terminus through the formation of an activated ester or thio ester
CC	intermediate. (I) has antiinflammatory activity and prevents accumulation
CC	of pancreatic digestive enzymes, and prevents exocytic fusion of vesicles
CC	containing secretory enzymes of pancreas. (I) is useful for treating
CC	acute pancreatitis. The present sequence represents the Clostridium
CC	botulinum BOMT/A neurotoxin heavy chain protoxin which is given in
CC	the exemplification of the present invention.
CC	
SQ	Sequence 423 AA;
Query Match	100.0%; Score 1071; DB 21; Length 423;
Best Local Similarity	100.0%; Pred. No. 3.4e-99;
Matches 207; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 IKVNWDFPSSSENFNDLNKGEITSDNTIEAAENISIDLIOQYYLTFFNDEPEN 60 DB 7 IKVNWMDFPSSSENFNDLNKGEITSNTIEAAEENSIDLLIQYYLTFNFDNEPEN 66 QY 61 ISENTSSDIIGQLMPLNIEREFPNGKKYEELDKYMFMHFLRAOEPFHGSRALATNSVE 120 DB 67 ISIENTSSDIIGQLMPLNIEREFPNGKKYEELDKYTMFHLRAOEHEGKSRLATNSVE 126 QY 121 ALNLSRYTYFTSSDYVKVKNKATEAAMFLGWEOLVYDFTETSETSVTTDKIADTTTT 180 DB 127 ALLNLSRYTYFTSSDYVKVKNKATEAAMFLGWEOVLVDFTDTETSETSVTTDKIADTTTT 186 QY 181 PYIGPALNIGMWLYKDDFVGALIIFSGA 207 DB 187 PYIGPALNIGMWLYKDDFVGALIIFSGA 213
RESULT 6	
LAY77140	

ID	AA77140	standard; Protein; 837 AA.
XX		
XX	AA77140;	
AC		
XX		
DT	08-MAY-2000	(first entry)
XX		
DE	Native botulinum neurotoxin serotype A (BONTA).	
XX		
XX	Botulinum neurotoxin; heavy chain; BONT; serotype A;	
RW	Venezuelan equine encephalitis virus replicon;	
KW	VEE; botulism; vaccine; diagnosis; drug screening.	
KW		
OS	Clostridium botulinum.	
XX		
FT	Key	Location/Qualifiers
FT	Misc-difference	837
XX	/note= "Apparently encoded by GGATGGGAG AAGGCCACT G"	
XX	WO200002524-A2.	
XX		
PD	20-JAN-2000.	
XX		
XX	09-JUL-1999;	99WO-US15570.
PE		
PR	10-JUL-1998;	98US-0092416.
PR	12-MAY-1999;	99US-0133870.
XX		
PA	(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.	
XX		
PI	Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;	
XX		
DR	WPI; 2000-160827/14.	
DR	N-PSDB; AAZ87218.	
XX		
PT	Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum	
PT	toxin serotypes A-G, is used for inducing an immune response against	
XX	botulinum -	
XX		
XX	Example 3; Page 49; 54pp; English.	
XX	<p>The invention relates to novel vaccines that induce a protective immune response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F and G (BONTA-BONTG). The vaccine of the invention is novel recombinant DNA construct comprising a vector, and at least one nucleic acid fragment comprising a C-terminal heavy chain fragment (Hc) from BONT serotypes A-G. In preferred embodiments of the invention, the vector is a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this vector results in the production of large amounts of a protein encoded by a sequence cloned into the replicon. The constructs are used to produce vaccines against botulism. The proteins can also be used as diagnostic tools for the diagnosis of botulism. The transformed host cells can be used to analyse the effectiveness of drugs and agents which inhibit toxin effects. The vaccine currently used against botulism is dangerous and expensive to produce, and contains formalin, which is very painful for the recipient. Also, the vaccine is incomplete, in that only 5 of the 7 serotypes are represented in the formulation. The novel vaccine of overcomes these problems, as it is easily purified, and available in large quantities. It is also expressed in the lymph nodes for a better immune response. The present sequence represents the native BONTA heavy chain used in an exemplification of the present invention.</p>	
XX		
XX	Sequence	837 AA;
XX		
XX	Query Match	100.0%; Score 1071; DB 21; Length 837;
XX	Best Local Similarity	100.0%; Pid. No. 8.9e-99;
XX	Matches 207; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
QY	1	IKVNNWDLFFSPSDNFPNDLNKGGELTSDPNIIEAAENISDLIQOXYITFPNPEPEN 60
QY		
QY	2	IKVNNWDLFFSPSDNFPNDLNKGGELTSDPNIIEAAENISDLIQOXYITFPNPEPEN 61
QY		
QY		ISINLSSDIIGOLELMPNIEFPGKKYELDKYTMFHYLRAOEFHCKSRITLNSVNE 120
QY		

Db 62 ISIENLSSDIIGOLELMPNIEERFPNGKKEYELDKYTMFHYLRAQEFHCKSRIALTNSVNE 121
 QY 121 ALLNSRYTFFSSDYVKVKNKATEAMFLGWEDLVYDFDETSEVSTTDKIADITITII 180
 Db 122 ALLNSRYTFFSSDYVKVKNKATEAMFLGWEDLVYDFDETSEVSTTDKIADITITII 181
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
 Db 182 PYIGPALNIGNMLYKDDFVGALIFSGA 208

RESULT 7
 ABG69075
 ID ABG69075 standard; Protein; 861 AA.
 XX
 AC ABG69075;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Botulinum neurotoxin light chain polypeptide #9a.
 XX
 KW Botulinum neurotoxin light chain; BoNT LC; botulism; dystonia; pain;
 KW spasticity; ocular motility; facial dyskinesia; stiff-person syndrome;
 KW bladder dysfunction; segmental myoclonus; hyperkinetic disorder;
 KW cosmetic treatment; facial wrinkle; cerebral palsy; analgesic; relaxant;
 KW lower motor neuron hyperactivity; autonomic nerve function; muscular;
 KW immunostimulant; antibacterial.
 XX
 OS Clostridium botulinum.
 XX
 PN W0200236758-A2.
 PD 10-MAY-2002.
 XX
 PF 06-NOV-2001; 2001WO-US47230.
 XX
 PR 06-NOV-2000; 2000US-246774P.
 PR 20-JUL-2001; 2001US-0910186.
 PR 09-AUG-2001; 2001US-311966P.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 PI Smith LA, Jensen M;
 XX
 DR WPI: 2002-575192/61.
 DR N-PSDB; ABR98545.
 XX
 PT Novel nucleic acid molecule encoding botulinum neurotoxin light chain
 PT serotype A, useful for producing the neurotoxin for vaccination against
 PT botulism, comprises sequence expressible in host other than Clostridium
 PT
 XX
 XX Claim 13; Page 133-134; 166pp; English.
 XX
 CC The invention relates to a nucleic acid molecule encoding a botulinum
 CC neurotoxin light chain (BoNT LC) serotype A, where the DNA has a sequence
 CC that is expressible in a host organism other than Clostridium, or has a
 CC total A+T content that is less than about 70%. The BoNT LC protein is
 CC useful in vaccination against botulism, for eliciting protective immunity
 CC in a mammal, for treating dystonias, spasticity, pain, ocular motility,
 CC facial dyskinesias, stiff-person syndrome, bladder dysfunction, segmental
 CC myoclonus, hyperkinetic disorders, cosmetic treatment of facial wrinkles,
 CC conditions characterised by hyperactivity of the lower motor neuron, and
 CC to control autonomic nerve function or tip-toe-walking due to stiff
 CC muscles common in children with cerebral palsy. The sequences are also
 CC useful for screening for botulinum neurotoxin inhibitors. This sequence
 CC represents a botulinum neurotoxin light chain serotype A protein.
 CC
 XX
 SO Sequence 861 AA;
 Query Match 100.0%; Score 1071; DB 23; Length 861;
 Best Local Similarity 100.0%; Pred. No. 9.3e-99;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMWDLFFSPSEDNFTNDLNKGEITSDTYIEAAEENISDLIQOYLYTFNPNNEPN 60
 Db 456 IKVNMWDLFFSPSEDNFTNDLNKGEITSDTYIEAAEENISDLIQOYLYTFNPNNEPN 515
 QY 61 ISIENLSSDIIGOLELMPNIEERFPNGKKEYELDKYTMFHYLRAQEFHCKSRIALTNSVNE 120
 Db 516 ISIENLSSDIIGOLELMPNIEERFPNGKKEYELDKYTMFHYLRAQEFHCKSRIALTNSVNE 575
 QY 121 ALLNSRYTFFSSDYVKVKNKATEAMFLGWEDLVYDFDETSEVSTTDKIADITITII 180
 Db 576 ALLNSRYTFFSSDYVKVKNKATEAMFLGWEDLVYDFDETSEVSTTDKIADITITII 635
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
 Db 636 PYIGPALNIGNMLYKDDFVGALIFSGA 662

RESULT 8
 AAM56019
 ID AAM56019 standard; Protein; 871 AA.
 XX
 AC AAM56019;
 XX
 DT 27-JUL-1998 (first entry)
 XX
 DE Recombinant botulinum neurotoxin type A Lm423/A (Q2E,N26K,A27Y).
 XX
 KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
 KW immunogen; detection; tetanus; non-toxic; toxin.
 XX
 OS Synthetic.
 OS Clostridium botulinum.
 XX
 PN W09807864-A1.
 PD 26-FEB-1998.
 XX
 PF 22-AUG-1997; 97WO-GB02273.
 XX
 PR 13-DEC-1996; 96GB-0025996.
 PR 23-AUG-1996; 96GB-0011671.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PA (SPEY-) SPEYWOOD LAB LTD.
 XX
 PI Foster KA, Quinn CP, Shone CC;
 XX
 DR WPI: 1998-169168/15.
 DR N-PSDB; AAV26291.
 XX
 PT Recombinant neurotoxin polypeptides - used to develop therapeutic
 PT agents, immunogens or as non-toxic standards for the detection of
 PT neurotoxins
 PT
 XX
 XX Example 1; Page 108-111; 137pp; English.
 XX
 CC The present sequence represents a recombinant neurotoxin protein from
 CC the present invention. The present invention describes recombinant
 CC neurotoxin proteins which comprise a first and second domain, where
 CC the first domain is adapted to cleave one or more vesicle or
 CC plasma-membrane associated proteins essential to exocytosis, and where
 CC the second domain is adapted: (a) to translocate the protein into a
 CC cell; (b) to increase the solubility of the protein compared to the
 CC solubility of the first domain on its own, or (c) both to translocate
 CC the protein into a cell and to increase the solubility of the protein
 CC compared to the solubility of the first domain on its own, the protein
 CC being free of clostridial neurotoxin (CN) and free of CN precursor that
 CC can be converted into toxin by proteolytic action. The recombinant
 CC proteins can be used as therapeutic agents for targeting cells
 CC expressing a relevant substrate. The products can also be used as
 CC immunogens and as non-toxic standards for the assessment and development
 CC of in vitro assays for the detection of functional botulinum or tetanus

CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 871 AA;
Query Match 100.0%; Score 1071; DB 19; Length 871;
Best Local Similarity 100.0%; Pred. No. 9.4e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IKVNMNDLFPSPEQNDFTNDLNGEETSDPTNEAAEENISLDLIQOYYLTFFNDNEPEN 60
DB 455 IKVNMNDLFPSPEQNDFTNDLNGEETSDPTNEAAEENISLDLIQOYYLTFFNDNEPEN 514
QY 61 ISEINSSDIIGOLELMPNIERFPNGKKYELDKYTFMFLRAQEFHGKSRIALTNSVNE 120
DB 515 ISEINSSDIIGOLELMPNIERFPNGKKYELDKYTFMFLRAQEFHGKSRIALTNSVNE 574
QY 121 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVEOLVYDFTDETSEVSTTDKADITITII 180
DB 575 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVEOLVYDFTDETSEVSTTDKADITITII 634
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 635 PYIGPALNIGNMLYKDDFVGALIFSGA 661
RESULT 9
AAM56007
ID AAM56007 standard; protein; 871 AA.
AC AAM56007;
XX
XX 27-JUL-1998 (first entry)
DE Recombinant botulinum neurotoxin type A LH423/A.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KM Immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.
XX
XX Clostridium botulinum.
XX
XX WO9807864-A1.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-GB02273.
XX
XX 13-DEC-1996; 96GB-0025996.
XX
XX 23-AUG-1996; 96GB-0017671.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX (SPEY-) SPEYWOOD LAB LTD.
XX
XX Foster KA, Quinn CP, Shone CC;
XX
XX WPI; 1998-169168/15.
XX
XX N-PsDB; AAV26279.
XX
XX Recombinant neurotoxin polypeptides - used to develop therapeutic
XX agents, immunogens or as non-toxic standards for the detection of
XX neurotoxins
XX
XX Example 1; Page 33-35; 137pp; English.
XX
XX The present sequence represents a recombinant neurotoxin protein from
XX the present invention. The present invention describes recombinant
XX neurotoxin proteins which comprise a first and second domain, where
XX the first domain is adapted to cleave one or more vesicle or
XX plasma-membrane associated proteins essential to exocytosis, and where
XX the second domain is adapted: (a) to translocate the protein into a
XX cell; (b) to increase the solubility of the protein compared to the
XX solubility of the first domain on its own, or (c) both to translocate
XX the protein into a cell and to increase the solubility of the protein

CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 871 AA;
Query Match 100.0%; Score 1071; DB 19; Length 871;
Best Local Similarity 100.0%; Pred. No. 9.4e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IKVNMNDLFPSPEQNDFTNDLNGEETSDPTNEAAEENISLDLIQOYYLTFFNDNEPEN 60
DB 455 IKVNMNDLFPSPEQNDFTNDLNGEETSDPTNEAAEENISLDLIQOYYLTFFNDNEPEN 514
QY 61 ISEINSSDIIGOLELMPNIERFPNGKKYELDKYTFMFLRAQEFHGKSRIALTNSVNE 120
DB 515 ISEINSSDIIGOLELMPNIERFPNGKKYELDKYTFMFLRAQEFHGKSRIALTNSVNE 574
QY 121 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVEOLVYDFTDETSEVSTTDKADITITII 180
DB 575 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVEOLVYDFTDETSEVSTTDKADITITII 634
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 635 PYIGPALNIGNMLYKDDFVGALIFSGA 661
RESULT 10
AAM56008
ID AAM56008 standard; protein; 871 AA.
AC AAM56008;
XX
XX 27-JUL-1998 (first entry)
DE Botulinum neurotoxin type A BONT/A.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KM Immunogen; detection; tetanus; non-toxic; toxin.
XX
XX Clostridium botulinum.
XX
XX WO9807864-A1.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-GB02273.
XX
XX 13-DEC-1996; 96GB-0025996.
XX
XX 23-AUG-1996; 96GB-0017671.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX (SPEY-) SPEYWOOD LAB LTD.
XX
XX Foster KA, Quinn CP, Shone CC;
XX
XX WPI; 1998-169168/15.
XX
XX N-PsDB; AAV26280.
XX
XX Recombinant neurotoxin polypeptides - used to develop therapeutic
XX agents, immunogens or as non-toxic standards for the detection of
XX neurotoxins
XX
XX Disclosure; Page 52-54; 137pp; English.
XX
XX The present sequence represents botulinum neurotoxin type A from
XX the present invention. The present invention describes recombinant
XX neurotoxin proteins which comprise a first and second domain, where

CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own; or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own; the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.

XX Sequence 871 AA;

Query Match 100.0%; Score 1071; DB 19; Length 871;

Best Local Similarity 100.0%; Pred. No. 9.4e-99;

Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNNWDLFFSPSDNFTNDLNKGEITSDTNIEAEBNISLDLIQOYYLTFFNDNEPEN 60

Db 455 IKVNNWDLFFSPSDNFTNDLNKGEITSDTNIEAEBNISLDLIQOYYLTFFNDNEPEN 514

QY 61 ISINLSSDIIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOFEHGKSRIALTNSVNE 120

Db 515 ISINLSSDIIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOFEHGKSRIALTNSVNE 574

QY 121 ALNPSRYVTFEFSDDYKKVKKATEAAMFLGWVOLYDFDETSEVSTDKADIITII 180

Db 575 ALNPSRYVTFEFSDDYKKVKKATEAAMFLGWVOLYDFDETSEVSTDKADIITII 634

QY 181 PYIGPALNIGMMLYKDDFVGALIFSGA 207

Db 635 PYIGPALNIGMMLYKDDFVGALIFSGA 661

Db 635 PYIGPALNIGMMLYKDDFVGALIFSGA 661

RESULT 11

AAW56016

ID AAW56016 standard; Protein; 873 AA.

AC AAW56016;

DT 27-JUL-1998 (first entry)

DE Recombinant botulinum neurotoxin type A 2LH423/A (Q2E,N26K,A27Y).

XX Botulinum; recombinant; Clostridium botulinum; neurotoxin;

KW immunogen; detection; tetanus; non-toxic; toxin.

XX Synthetic.

OS Clostridium botulinum.

PN WO9807864-A1.

PD 26-FEB-1998.

PF 22-AUG-1997; 97WO-GB02273.

OS 13-DEC-1996; 96GB-0025996.

PR 23-AUG-1996; 96GB-0017671.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PA (SPEY-) SPEYWOOD LAB LTD.

PI Foster KA, Quinn CP, Shone CC;

XX WPI: 1998-169168/15.

DR N-PSDB; AAV26288.

XX Recombinant neurotoxin polypeptides - used to develop therapeutic

PT agents, immunogens or as non-toxic standards for the detection of

PT neurotoxins

XX Example 1; Page 45-48; 137pp; English.

XX The present sequence represents a recombinant neurotoxin protein from

CC the present invention. The present invention describes recombinant

CC neurotoxin proteins which comprise a first and second domain, where

CC the first domain is adapted to cleave one or more vesicle or

CC plasma-membrane associated proteins essential to exocytosis, and where

CC the second domain is adapted: (a) to translocate the protein into a

CC cell; (b) to increase the solubility of the protein compared to the

CC solubility of the first domain on its own; or (c) both to translocate

CC the protein into a cell and to increase the solubility of the protein

CC compared to the solubility of the first domain on its own; the protein

CC being free of clostridial neurotoxin (CN) and free of CN precursor that

CC can be converted into toxin by proteolytic action. The recombinant

CC proteins can be used as therapeutic agents for targeting cells

CC expressing a relevant substrate. The products can also be used as

CC immunogens and as non-toxic standards for the assessment and development

CC of in vitro assays for the detection of functional botulinum or tetanus

CC neurotoxins either in foodstuffs or in environmental samples.

XX Sequence 873 AA;

Query Match 100.0%; Score 1071; DB 19; Length 873;

Best Local Similarity 100.0%; Pred. No. 9.5e-99;

Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNNWDLFFSPSDNFTNDLNKGEITSDTNIEAEBNISLDLIQOYYLTFFNDNEPEN 60

Db 457 IKVNNWDLFFSPSDNFTNDLNKGEITSDTNIEAEBNISLDLIQOYYLTFFNDNEPEN 516

QY 61 ISINLSSDIIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOFEHGKSRIALTNSVNE 120

Db 517 ISINLSSDIIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOFEHGKSRIALTNSVNE 576

QY 121 ALNPSRYVTFEFSDDYKKVKKATEAAMFLGWVOLYDFDETSEVSTDKADIITII 180

Db 577 ALNPSRYVTFEFSDDYKKVKKATEAAMFLGWVOLYDFDETSEVSTDKADIITII 636

QY 181 PYIGPALNIGMMLYKDDFVGALIFSGA 207

Db 637 PYIGPALNIGMMLYKDDFVGALIFSGA 663

Db 637 PYIGPALNIGMMLYKDDFVGALIFSGA 663

RESULT 12

AAW56009

ID AAW56009 standard; Protein; 875 AA.

AC AAW56009;

DT 27-JUL-1998 (first entry)

DE Recombinant botulinum neurotoxin type A L/4H423/A.

XX Botulinum; recombinant; Clostridium botulinum; neurotoxin;

KW immunogen; detection; tetanus; non-toxic; toxin.

XX Synthetic.

OS Clostridium botulinum.

PN WO9807864-A1.

PD 26-FEB-1998.

PF 22-AUG-1997; 97WO-GB02273.

OS 13-DEC-1996; 96GB-0025996.

PR 23-AUG-1996; 96GB-0017671.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PA (SPEY-) SPEYWOOD LAB LTD.

PI Foster KA, Quinn CP, Shone CC;
XX
XX WPI: 1998-169168/15.
DR N-PSDB: AAV26281.
XX
PT Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents; immunogens or as non-toxic standards for the detection of
PT neurotoxins
XX
XX
PS Example 1: Page 58-60; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 875 AA:

Query Match 100.0%; Score 1071; DB 19; Length 875;
Best Local Similarity 100.0%; Pred. No. 9.5e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMWDLFFSSSEDNFTNDLKNKGEITSPTNIEAAENISDLIQOYYLTFNFPNEPEN 60
DB 459 IKVNMWDLFFSSSEDNFTNDLKNKGEITSPTNIEAAENISDLIQOYYLTFNFPNEPEN 518
QY 61 ISTEINSSDIIGOLELMPNIEERFPNGKYEIDKYMFMHFLRAQEFHGSRIALNSVNE 120
DB 519 ISTEINSSDIIGOLELMPNIEERFPNGKYEIDKYMFMHFLRAQEFHGSRIALNSVNE 578
QY 121 ALLNPSRVYTFSSDYVKKVKNKATEAAMFLGWVEQLVYDFDETSEVSTTDKIADITITII 180
DB 579 ALLNPSRVYTFSSDYVKKVKNKATEAAMFLGWVEQLVYDFDETSEVSTTDKIADITITII 638
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 639 PYIGPALNIGNMLYKDDFVGALIFSGA 665

RESULT 13
AAW56010
ID AAW56010 standard; Protein: 878 AA.
XX
AC AAW56010;
XX
DT 27-JUL-1998 (first entry)
XX
DE Recombinant botulinum neurotoxin type A LFXA/3H423/A.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN WO9807864-A1.
XX
XX 26-FEB-1998.
PD
XX 22-AUG-1997; 97WO-GB02273.
PF

XX
PR 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
XX
XX Foster KA, Quinn CP, Shone CC;
XX
XX
DR WPI: 1998-169168/15.
DR N-PSDB: AAV26282.
XX
XX
PS Example 1: Page 64-66; 137pp; English.
XX
CC The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
XX
SQ Sequence 878 AA:

Query Match 100.0%; Score 1071; DB 19; Length 878;
Best Local Similarity 100.0%; Pred. No. 9.5e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMWDLFFSSSEDNFTNDLKNKGEITSPTNIEAAENISDLIQOYYLTFNFPNEPEN 60
DB 462 IKVNMWDLFFSSSEDNFTNDLKNKGEITSPTNIEAAENISDLIQOYYLTFNFPNEPEN 521
QY 61 ISTEINSSDIIGOLELMPNIEERFPNGKYEIDKYMFMHFLRAQEFHGSRIALNSVNE 120
DB 522 ISTEINSSDIIGOLELMPNIEERFPNGKYEIDKYMFMHFLRAQEFHGSRIALNSVNE 581
QY 121 ALLNPSRVYTFSSDYVKKVKNKATEAAMFLGWVEQLVYDFDETSEVSTTDKIADITITII 180
DB 582 ALLNPSRVYTFSSDYVKKVKNKATEAAMFLGWVEQLVYDFDETSEVSTTDKIADITITII 641
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 642 PYIGPALNIGNMLYKDDFVGALIFSGA 668

RESULT 14
AAW56015
ID AAW56015 standard; Protein: 894 AA.
XX
AC AAW56015;
XX
DT 27-JUL-1998 (first entry)
XX
DE Recombinant botulinum neurotoxin type A 23LH423/A (Q2E, N26K, A27Y).
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW immunogen; detection; tetanus; non-toxic; toxin.
XX
OS Synthetic.

OS Clostridium botulinum.
XX
PN WO9807864-A1.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-GB02273.
XX
PR 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
PI Foster KA, Quinn CP, Shone CC;
XX
DR WPI; 1998-169168/15.
DR N-PSDB; AAV26287.
XX
XX Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
PS
PS Example 1; Page 39-42; 137pp; English.
XX
XX The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
CC
SQ Sequence 894 AA;
XX
Query Match 100.0%; Score 1071; DB 19; Length 894;
Best Local Similarity 100.0%; Pred. No. 9.8e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IKVNMWDLFFSPSEDNFNNDLKGGEITSPTNIEAEENISIDLIOQYVLPFNPNNEPN 60
DB 478 IKVNMWDLFFSPSEDNFNNDLKGGEITSPTNIEAEENISIDLIOQYVLPFNPNNEPN 537
QY 61 ISENLSSDIIIGOLELMPNIEERFPNGKRYELDKYTMFHYLRAQEFHKSRIALTNSVNE 120
DB 538 ISENLSSDIIIGOLELMPNIEERFPNGKRYELDKYTMFHYLRAQEFHKSRIALTNSVNE 597
QY 121 ALLNRSRYTFFSSDYVKVKVKAATPAAMFLGWEOVLVDFDETSEVSTTKIADITITII 180
DB 598 ALLNRSRYTFFSSDYVKVKVKAATPAAMFLGWEOVLVDFDETSEVSTTKIADITITII 657
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 658 PYIGPALNIGNMLYKDDFVGALIFSGA 684
RESULT 15
ID AAW56012
XX AAW56012 standard; Protein: 907 AA.
AC AAW56012;
XX
DT 27-JUL-1998 (first entry)

XX
DE Recombinant botulinum neurotoxin type A LfXa/3H423/A-CtxA 14.
XX
KW Botulinum; recombinant; Clostridium botulinum; neurotoxin;
KW Immunogen; detection; tetanus; non-toxic; toxin.
OS Synthetic.
OS Clostridium botulinum.
XX
XX WO9807864-A1.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-GB02273.
XX
XX 13-DEC-1996; 96GB-0025996.
PR 23-AUG-1996; 96GB-0017671.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA (SPEY-) SPEYWOOD LAB LTD.
XX
PI Foster KA, Quinn CP, Shone CC;
XX
XX WPI; 1998-169168/15.
DR N-PSDB; AAV26284.
XX
XX Recombinant neurotoxin polypeptides - used to develop therapeutic
PT agents, immunogens or as non-toxic standards for the detection of
PT neurotoxins
PS
PS Example 1; Page 77-79; 137pp; English.
XX
XX The present sequence represents a recombinant neurotoxin protein from
CC the present invention. The present invention describes recombinant
CC neurotoxin proteins which comprise a first and second domain, where
CC the first domain is adapted to cleave one or more vesicle or
CC plasma-membrane associated proteins essential to exocytosis, and where
CC the second domain is adapted: (a) to translocate the protein into a
CC cell; (b) to increase the solubility of the protein compared to the
CC solubility of the first domain on its own, or (c) both to translocate
CC the protein into a cell and to increase the solubility of the protein
CC compared to the solubility of the first domain on its own, the protein
CC being free of clostridial neurotoxin (CN) and free of CN precursor that
CC can be converted into toxin by proteolytic action. The recombinant
CC proteins can be used as therapeutic agents for targeting cells
CC expressing a relevant substrate. The products can also be used as
CC immunogens and as non-toxic standards for the assessment and development
CC of in vitro assays for the detection of functional botulinum or tetanus
CC neurotoxins either in foodstuffs or in environmental samples.
CC
SQ Sequence 907 AA;
XX
Query Match 100.0%; Score 1071; DB 19; Length 907;
Best Local Similarity 100.0%; Pred. No. 1e-98;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IKVNMWDLFFSPSEDNFNNDLKGGEITSPTNIEAEENISIDLIOQYVLPFNPNNEPN 60
DB 462 IKVNMWDLFFSPSEDNFNNDLKGGEITSPTNIEAEENISIDLIOQYVLPFNPNNEPN 521
QY 61 ISENLSSDIIIGOLELMPNIEERFPNGKRYELDKYTMFHYLRAQEFHKSRIALTNSVNE 120
DB 522 ISENLSSDIIIGOLELMPNIEERFPNGKRYELDKYTMFHYLRAQEFHKSRIALTNSVNE 581
QY 121 ALLNRSRYTFFSSDYVKVKVKAATPAAMFLGWEOVLVDFDETSEVSTTKIADITITII 180
DB 582 ALLNRSRYTFFSSDYVKVKVKAATPAAMFLGWEOVLVDFDETSEVSTTKIADITITII 641
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
DB 642 PYIGPALNIGNMLYKDDFVGALIFSGA 668

Search completed: March 13, 2003, 11:39:09
Job time : 41.7752 secs

R:DasGupta, B.R.; Dekleva, M.L.
 Biochimie 72, 661-664, 1990
 A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and aro
 A:Reference number: A60025; MUID:91120847; PMID:2126206
 A:Accession: A60025
 A:Molecule type: protein
 A:Residues: 2-6:445-453, 'x', 455-457 <DASL>
 R:DasGupta, B.R.; Foley, J.; Niece, R.
 Biochemistry 26, 4162, 1987
 A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.
 A:Reference number: A27000
 A:Accession: A27000
 A:Molecule type: protein
 A:Residues: 2-47 <DAS2>
 R:Bliz, T.; Blasi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
 J. Biol. Chem. 269, 1617-1620, 1994
 A:Title: Proteolysis of SNAP-25 by types E and A botulinial neurotoxins.
 A:Reference number: A49708; MUID:94124495; PMID:8294407
 A:Contents: annotation
 A:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
 C:Genetics:
 A:Gene: atx; botA
 C:Function:
 A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate
 C:Superfamily: tetanus toxin
 C:Keywords: disulfide bond; hydrolase; metalloprotease; neurotoxin; transmembrane prot
 F:2-444/Product: bontoxylisin A light chain #status experimental <LGHT>
 F:445-1296/Product: bontoxylisin A heavy chain #status experimental <HVT>
 F:223,227/Binding site: zinc (His) #status predicted
 F:224/Active site: Glu #status predicted

Query Match 100.0%; Score 1071; DB 1; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 1.3e-77;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 60
 |||||
 DB 455 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 514
 QY 61 ISTEENSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOEFHGKSRILATNSVNE 120
 |||||
 DB 515 ISTEENSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOEFHGKSRILATNSVNE 574
 QY 121 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVQOLYVDFDTSEVSTDKIADITITII 180
 |||||
 DB 575 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVQOLYVDFDTSEVSTDKIADITITII 634
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
 |||||
 DB 635 PYIGPALNIGNMLYKDDFVGALIFSGA 661

RESULT 2
 140645
 botulinum neurotoxin type A - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
 C:Accession: 140645
 R:Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
 Res. Microbiol. 144, 547-556, 1993
 A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
 A:Reference number: 140645; MUID:94143603; PMID:8310180
 A:Accession: 140645
 A:status: preliminary
 A:Molecule type: translated from GB/EMBL/DBJ
 A:Residues: 1-1296 <RES>
 A:Cross-references: EMBL:X73423; NID:g507070; PIDN:CAAS1824.1; PID:g507071
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 85.7%; Score 918; DB 2; Length 1296;
 Best Local Similarity 85.4%; Pred. No. 2.4e-65;
 Matches 176; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 1 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 60
 |||||
 DB 455 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 514
 QY 61 ISTEENSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOEFHGKSRILATNSVNE 120
 |||||
 DB 515 ISTEENSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAOEFHGKSRILATNSVNE 574
 QY 121 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVQOLYVDFDTSEVSTDKIADITITII 180
 |||||
 DB 575 ALINPSRVYTFESSDYKKVKNKATEAMFLGWVQOLYVDFDTSEVSTDKIADITITII 634
 QY 181 PYIGPALNIGNMLYKDDFVGALIFSG 206
 |||||
 DB 635 PYIGPALNIGNMLYKDDFVGALIFSG 660

RESULT 3
 533411
 botulinum neurotoxin type F - Clostridium baratii
 C:Species: Clostridium baratii
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C:Accession: 533411; S31860
 R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P
 FEMS Microbiol. Lett. 108, 175-182, 1993
 A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neuroto
 A:Reference number: 533411; MUID:93252228; PMID:8486245
 A:Accession: 533411
 A:status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1268 <THO>
 A:Cross-references: EMBL:X68262; NID:9449138; PIDN:CAA48329.1; PID:949139
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 38.0%; Score 407.5; DB 2; Length 1268;
 Best Local Similarity 44.0%; Pred. No. 1.8e-24;
 Matches 95; Conservative 39; Mismatches 61; Indels 21; Gaps 7;

QY 1 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 59
 |||||
 DB 436 IKVNMNDFEPPSEDNFTNDLKGEEITSDTNEAAEENISLDLQOYYLTFFNDPEN 487
 QY 60 NISIEENSSDIIGOL---ELMPNIEFPNG---KKYELDKYTMFHYLRAOEFHGKSR 111
 |||||
 DB 488 --ALPNLSRLNLTQAQNSVYPKYD--SNGTSEIKEYTVDKLNFVYLAQKAPGESA 543
 QY 112 IATNSVNEALLNPSRVYTFESSDYKKVKNKATEAMFLGWVQOLYVDFDTSEVSTDD 171
 |||||
 DB 544 ISLTSSVNTALLDASKVYTFESSDYKKTVPVQALFTSWIQOYLNDFTTATQKSTID 603
 QY 172 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 207
 |||||
 DB 604 KIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGA 639

RESULT 4
 539791
 neurotoxin - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
 C:Accession: 539791
 R:Campbell, K.; Collins, M.D.; East, A.K.
 Biochim. Biophys. Acta 1216, 487-491, 1993
 A:Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridiu
 A:Reference number: 539791; MUID:94092745; PMID:8268233
 A:Accession: 539791
 A:status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1297 <CAM>
 A:Cross-references: EMBL:X74162; NID:9441275; PIDN:CAAS2275.1; PID:9441276
 C:Superfamily: tetanus toxin

C;Keywords: neurotoxin

Query Match	35.2%;	Score 377;	DB 2;	Length 1297;
Best Local Similarity	41.3%;	Pred. No. 5.2e-22;		
Matches 88; Conservative	34;	Mismatches 75;	Indels 16;	Gaps 3;

QY	1	IKVNNMDFFSFESENFNDLNCGEITSDNIEAEENISLD-1IQOYLYTFNPNRPE	59
		: : : : : : : : : : : : : : :	
Db	451	IIVNNEDLEFANKDQSKDLAKAETIAYNTQNNITENNESIDQILINDLSSGIDLPNE	510
		: : : : : : : : : : : : : : : : : : : :	
QY	60	NISIENTSSDI-----IGQELMERNIERFPGKKYELDKYTMFHYLRAQEEFHGKSRIAL	114
		: : : : : : : : : : : : : : : : : : : :	
Db	511	NTEPFTNRDDIDIPVIKQ\$AL-----KRIYVGDSLFELYLAQTFRPSNIENIQL	560
		: : : : : : : : : : : : : : : : : : : :	
QY	115	TNSVNEALLNPSRYITFFSSDYVKKVKNKATEAMFLGVOLVYDFDETSEVSTTDKIA	174
		: : : : : : : : : : : : : : : : : : : :	
Db	561	TNSLNDALRNKKKYTFEFTSLVYERKANTVVGASLFVNNVKGVIDFTSE\$TKSTIDKVS	620
		: : : : : : : : : : : : : : : : : : : :	
QY	175	DITIIPIYGALNIGNMLYKDDFVGALIESGA	207
		: : : : : : : : : : : : : : : : : : : :	
Db	621	DVSIITPIYGALNVGN\$ETAKENFKNAEVEIGA	653
		: : : : : : : : : : : : : : : : : : : :	

RESULT 5

botulinum neurotoxin type E precursor - Clostridium butyricum
C/Species: Clostridium butyricum
C/Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C/Accession: JH0256; S16145
R/Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A/Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type E
A/Reference number: JH0256; MUID:92181428; PMID:1543481
A/Accession: JH0256
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-27, 'E', 29-1251 <PDU>
A/Cross-references: EMBL:X62088; NID:940379
A/Experimental source: strains ATCC 43181 and ATCC 43755
J Gen. Microbiol. 137, 519-525, 1991
J Gen. Microbiol. 137, 519-525, 1991
A/Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin
A/Reference number: S16145; MUID:9123315; PMID:2053376
A/Accession: S16145
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-229, 'M', 231-252 <FNU>
A/Cross-references: EMBL:X53180; NID:940407; PIDN:CA37321.1; PID:940408
A/Experimental source: Strain BL6340
C/Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C/Comment: The heavy chain mediates the binding of toxin to cell receptors while the light
C/Keyfamily: tetanus toxin
C/Keywords: neurotoxin
F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F:423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F:412-426/Dissulfide bonds: #status predicted

Query Match	35.0%;	Score 374.5;	DB 2;	Length 1251;
Best Local Similarity	38.9%;	Pred. No. 7.9e-22;		
Matches	84;	Conservative	51;	Mismatches 60;
			Indels	21;
			Gaps	7;

[illegible]

Qy 172 KIADITIIIPYIGPALNIGNMLKYKDFVGLIIFSGA 2077
 ||||:|:||||| | :| |
 Db 595 KIADISIVPYIGLALNIGNEAQKGNKDALELIGA 6300

RESULT 6

botulinum neurotoxin type E precursor - Clostridium botulinum

C:Species: Clostridium botulinum

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999

C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111

R:Whelan, S.M.; Elmore, M.D.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.

Eur. J. Biochem. 204, 657-667, 1992

A>Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin

A:Reference number: S21178; MUID:92174922; PMID:1541280

A:Accession: S21178

A:Molecule type: DNA

A:Residues: 1-1252 <WHE>

A:Cross-references: EMBL:X62683; NID:940397; PIDD:CAA44558.1; PID:940398

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A>Title: Gene probes for identification of the botulinum neurotoxin gene and strain

A:Reference number: S48107; MUID:94013372; PMID:8408542

A:Accession: S48107

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 616-982 <CAM>

A:Cross-references: EMBL:X70815; NID:9407786; PIDD:CAA50146.1; PID:9407787

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

R:Poulet, S.; Hauser, D.; Quanz, W.; Nlemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A>Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type E

A:Reference number: JH0257; MUID:92181428; PMID:153481

A:Accession: JH0257

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'T', 774-962, 'FE', 965-966, 'R', 96

A:Cross-references: EMBL:X62089; NID:940393; PIDD:CAA43999.1; PID:940394

A:Experimental source: Strain Beluga

R:Bhuz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Nlemann, H.

J. Biol. Chem. 265, 9153-9158, 1990

A>Title: The complete sequence of botulinum neurotoxin type A and comparison with other

A:Reference number: A35294; MUID:90264400; PMID:2160960

A:Accession: B35294

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-176, 'R', 178-252 <BIN>

A:Experimental source: Strain Beluga

R:Gimenez, J.A.; Dasgupta, B.R.

Biochimie 72, 213-217, 1990

A>Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals

A:Reference number: A60027; MUID:90344918; PMID:2116911

A:Accession: A60027

A:Molecule type: protein

A:Residues: 420-427 <GIM>

A:Experimental source: Strain Beluga

A>Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin

C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit

C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

E:1-422/product: botulinum neurotoxin type E light chain #status predicted <LCH>

E:423-1252/product: botulinum neurotoxin type E heavy chain #status predicted <HCH>

E:412-426/disulfide bonds: #status predicted

Query Match	34.4%	Score 368.5	DB 2	Length 1252
Best Local Similarity	38.48%	Pred. No. 2.4e-21		
Matches 83; Conservative	51;	Mismatches 61;	Indels 21;	Gaps 7;

```
Oy      1 IKVNNWDFEESPSEDNFTND-LNKGEET---TSDNTAEAAEFNISLDLIDQYYLTFFNF 55  
|::|| ||| |:::: :| :| ||:| ::||  
Db    427 IEINNGELFEFVASSENSYNDOKINTPKLEDDIYSNNNYE-----INDDQIVLANNSE 478
```

OY 56 NEPEISITSENISDIIQGLEMLPNIERPPNG----KKYEIDKYTFWHFYLRQAEFHGKSR 111
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 479 SAP-GLSDEKSLNTLTQN-D-AIIPKRD--SNCSLDIEQHADVLEINFFELDDAKQKEGGNN 534
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY 112 IALNSVNEALLNPSRVYTFSSDYVKRKNKATEAMFLGVWEDVLVDFTDETSESVSTTD 171
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 535 VNLSTSDIALTEOPKRITYTFESFSEFINNVNKPVQAALFPVMLOQLVLDFTFEANQKSTVD 594
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
OY 172 KIAADTITTIPTGPAINTGNMILYKDPFGALIFSGA 207
|||||||:||||| ||||| | : || || ||
Db 595 KIAIDSIIVPYIGLALNTIGNENQKGFKALELLGA 630

RESULT 7
A48940
botltoxllysin (EC 3.4.2.4.69) B precursor - Clostridium botulinum
N:Alternate names: botulinum neurotoxin type B (BoNT/B)
C:Species: Clostridium botulinum
C:date: 19-Dec-1993 #sequence revision 18-Nov-1994 #text change 18-Jun-1999
C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
R:Methan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
Appl. Environ. Microbiol. 58, 2345-2354, 1992
A:title: Molecular cloning of the Clostridium botulinum structural gene encoding the toxin
A:Reference number: A48940; MWID:92384550; PMID:1514783
A:Accession: A48940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1291 <MEM>
A:Cross-references: EMBL:X70817; NID:g144734; PIDN:AAA23211.1; PID:g144735
A:Experimental source: type B, Danish
A>Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIRP:112081); this publication
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:title: Gene probes for identification of the botulinum neurotoxin gene and specific identification
A:Reference number: S48103; MWID:94013372; PMID:8408542
A:Accession: S48105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 634-994 <CAM>
A:Cross-references: EMBL:Z11934; NID:9407782; PIDN:CAA50148.1; PID:g407783
A:Experimental source: proteolytic type B, strain NCTC 7273
R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A:description: Partial amino acid sequence of botulinum neurotoxin type B and comparison
A:Reference number: S21575
A:Accession: S21575
A:Molecule type: DNA
A:Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZ>
A:Cross-references: EMBL:Z11934; NID:940383; PIDN:CAAT7991.1; PID:g40384
R:Kurazono, H.; Mochida, S.; Blinz, T.; Etzel, U.; Quanz, M.; Grebenstein, O.; Wernars, K.
J. Biol. Chem. 267, 14721-14729, 1992
A:title: Minimal essential domains specifying toxicity of the light chains of tetanus toxin
A:Reference number: A42871; MWID:92340509; PMID:1634516
A:Accession: A42871
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-313, 'S', 315-451 <KUR>
A:Experimental source: strain OKra
A>Note: sequence extracted from NCBI backbone (NCBIRP:109365)
R:Dasgupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A:title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
A:Reference number: S07155; MWID:89000987; PMID:3139097
A:Accession: S07155
A:Molecule type: protein
A:Residues: 2-29, 'M', 31-45 <DAS>
A:Accession: S08562
A:Molecule type: protein
A:Residues: 442-463, 'R', 465-467 <DA>
R:Schmidt, J.J.; Sathymocorthy, V.; Dasgupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A:title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A:Reference number: S07128; MWID:85197963; PMID:3888113

A:Accession: S07128
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 2-16 <SCH1>
 A:Accession: S08573
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 2-17 <SCH2>
 A:Accession: S08574
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 442-459 <SCH3>
 R:Schiano, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; DasGupta, Nature 359, 832-835, 1992
 A:Title: Tetanus and Botulinum-B neurotoxins block neurotransmitter release by proteolysis
 A:Reference number: S27125; PMID:93063293; PMID:1331807
 A:Contents: annotation
 C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses
 C:Genetics:
 A:Gene: bont/b
 C:Function:
 A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
 C:Superfamily: tetanus toxin
 C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F:2-441/Product: bontoxilysin B light chain #status experimental <Lght>
 F:442-1291/Product: bontoxilysin B heavy chain #status experimental <Hvy>
 F:230-234/Binding site: zinc (His) #status predicted
 F:231/Active site: Glu #status predicted

Query Match

Best Local Similarity 39.4%; Pred. No. 3.6e-21;
Matches 87; Conservative 39; Mismatches 62; Indels 33; Gaps 4.

QY 1 IKVNNWDLFFSPSEDNFTNDLNKGEITSSDINIAEAENISLDLIQQYLLTFNFDNEPEN 60

```

Db 447 IDVDNEDFFIADKNSFSDLSKNERIEYNT-----QSNYIENDF---PIN 489

```

61 ISENSSDIIGOLEL-----MPNTERFPNGKYEELDKYTMFHYLRAQEE 106

Db 490 ELI--LDTDLISKIELPSENTESLTENDVNDVPRVYEKQPAIKKIFTDENIIFQYLXSQTFP 547

QY 107 HGKSRIALNSVNEALLNPISRVYTFPSSDYVKVKNKATEAAMELGWVEQLVYDPTDETSE 166

Db 548 LDIRDISLTSSFDALLFSNKKVYSFESMDYIKTKANKVEAGLEAGWVKQIVNDEVEIANK 6077

```
0Y      167 VSTTDKADITIIIPYIGPALINIGNMLKYDDFVGLIFESGA 207
       ::|||:::||::|||::|||::|||::|||::|||::|||::|||
```

Db 608 SNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFETIACA 648

RESULT 8

1A0631 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botuli

C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999

C:Accession::140631; S48103; S48104; S50015
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Date Modified 28 101-110 1004

ATitle: Nucleotide sequence of the gene coding for non-proteolytic Clostridium Microbium 26, 101-110, 1994
AReference number: 140631. MIMD.04132659. DMT.7764370

A:Accession: I40631
A:Reference Number: 140031, M01D.2412203, EMD.170430
A:Status: preliminary. translated from GR/EMBL/DDP.J

A:Residues: 1-1291 <RES>
A:Molecule type: DNA
A:Residues: 1-1291 <RES>

A: Cross-references: EMBL: X71343; NID: g296148; PIDN: CAA50482.1; PID: g296145
R: Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48103

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA

A:Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAML>

[illegible]

RESULT 14

botulinum neurotoxin type C1 precursor - Clostridium botulinum phage (type C, strain C-1)
C:Species: Clostridium botulinum phage
C:Date: 10-Mar-1994 #sequence.revision 07-Apr-1994 #text_change 23-Mar-2001
C:Accession: S11291, A35396, S22166, A49777
R:Hauser, D.; Eklund, M.W.; Kutazono, H.; Blinz, T.; Niemann, H.; Gill, D.M.; Boquet, P.;
Nucleic Acids Res. 18, 4924, 1990
A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
A:Reference number: S11291; MUID:90370487; PMID:2204031
A:Accession: S11291
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84, 'P', 86-1291 <RAN>
A:Cross-references: EMBL:X53751; NID:914905; PIDD:CAA37780.1; PID:914906
R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi, K.
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A:Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin
A:Reference number: A35396; MUID:91024998; PMID:2222445
A:Accession: A35396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-669, 'R', 671-1291 <TS1>
R:Tsuzuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
Submitted to the EMBL Data Library, December 1991
A:Disctiption: Nucleotide sequence of the gene for one of the components of hemagglutinin
A:Reference number: S22163
A:Accession: S22166
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1291 <TS2>
A:Cross-references: EMBL:X62389; NID:9558175; PIDD:CAA44263.1; PID:940390
R:Kimura, K.; Fujii, N.; Tsuzuki, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Oguma, K.
Appl. Environ. Microbiol. 57, 1166-1172, 1991
A:Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and whic
A:Reference number: A49777; MUID:91282468; PMID:2059039
A:Accession: A49777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-607 <TS3>
A:Cross-references: GB:D90210
C:Superfamily: tetanus toxin
;Keywords: neurotoxin

Query Match	24.6%	Score 263	DB 2	Length 1291
Best Local Similarity	31.2%	Pred. No.	7e-13	
Matches 64	Conservative 39	Mismatches 94	Indels 8	Gaps 4

[illegible]

QY 122 LINSRVYTFEFSDDYKVNKATAAEFLGVEOLVYDFETDESEVSTGTITIIIP 181
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 572 LDNSAKVYTYEPT-LANKVNAVGGEGLEFLMANVYVEDFTTNILKRTLDKISDVSAIIP 630

QY 182 YIGPALNIGNNMLYKDFEGVALIFSG 206
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 631 YIGPALNINISVNRKGFTEAAVNG 655

RESULT 15

hypothetical protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: A84977
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: A84977
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-246 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: yf10; BU402

Query Match	10.0%	Score 107;	DB 2;	Length 246;
Best Local Similarity	21.8%	Pred. No.	0.24;	
Matches	44;	Conservative	40;	Mismatches 44;
				Gaps 8;

[illegible]

Search completed: March 13, 2003, 11:41:44
Job time : 21.1066 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:47 ; Search time 31.0202 Seconds
(without alignments)
1374.968 Million cell updates/sec

Title: US-09-917-791-21
Perfect score: 1071
Sequence: 1 IKVNMWDLFFSPSEDFNTD.....NIGMWLYKDFVGLIFSGA 207

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB ID	Description
1	407.5	38.0	1268	2	Q45851 clostridium
2	374.5	35.0	1255	2	Q9FAR6 clostridium
3	368.5	34.4	1251	2	Q9K395 clostridium
4	367.5	34.3	1291	2	Q933K0 clostridium
5	366.5	34.2	1291	2	Q93G71 clostridium
6	366.5	34.0	1291	2	Q92A78 clostridium
7	364.5	34.0	1278	2	Q92A78 clostridium
8	357.5	33.4	1291	2	Q92A78 clostridium
9	347.5	32.4	1280	2	Q92A78 clostridium
10	333	31.1	173	2	Q9K88 clostridium
11	331	30.9	173	2	Q9K88 clostridium
12	330	30.8	173	2	Q9K88 clostridium
13	330	30.8	173	2	Q9K88 clostridium
14	314	29.3	1310	2	Q93N27 clostridium
15	303.5	28.3	1285	2	Q9LBR1 clostridium
16	303.5	28.3	1285	2	Q45967 clostridium

17	290.5	27.1	1275	12	Q90T67 clostridium
18	263	24.6	1280	2	Q45849 clostridium
19	263	24.6	1291	2	Q93HT3 clostridium
20	262	24.5	1280	2	Q9LBS7 clostridium
21	106.5	9.9	1197	2	Q45888 clostridium
22	106	9.9	1196	2	Q53550 clostridium
23	106	9.9	1196	2	Q45916 clostridium
24	106	9.9	1196	2	Q93HT4 clostridium
25	106	9.9	1196	9	Q92X77 clostridium
26	106	9.9	1196	9	Q92X77 clostridium
27	105	9.8	1196	5	Q9GSC4 clostridium
28	105	9.8	440	5	Q9GSC0 clostridium
29	105	9.8	440	5	Q9GSC0 clostridium
30	105	9.8	2771	5	Q9GSC8 clostridium
31	104	9.7	1193	2	P71108 clostridium
32	103	9.6	432	5	Q9GSC5 clostridium
33	103	9.6	432	5	Q9GSC1 clostridium
34	103	9.6	432	5	Q9GSC1 clostridium
35	103	9.6	1163	2	Q45850 clostridium
36	101	9.4	432	5	Q9GSC8 clostridium
37	101	9.4	432	5	Q9GSC7 clostridium
38	101	9.4	1198	2	O06018 clostridium
39	100.5	9.4	356	2	Q9AKM7 clostridium
40	100	9.3	664	5	O97263 clostridium
41	99.5	9.3	1055	12	O65146 african swi
42	99.5	9.3	2747	5	O9BJX9 clostridium
43	99	9.2	1193	2	O45914 clostridium
44	99	9.2	1193	2	P71107 clostridium
45	99	9.2	1196	2	Q9LBS8 clostridium

ALIGNMENTS

RESULT 1

Q45851 PRELIMINARY: PRT; 1268 AA.
AC Q45851:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Neurotoxin type F.
GN BONT /F.
OS Clostridium baratii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93252228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
RT neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEBS Microbiol. Lett. 108:175-182(1993).
DR EMBL: X68262; CAA48329.1; -
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -
DR InterPro: IPR000395; Bontolysin.
DR InterPro: IPR000130; Zn_MPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PD00760; BONTOLYSIN.
DR PRODOM: PD001963; Bontolysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SO SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 38.0%; Score 407.5; DB 2; Length 1268;

Best Local Similarity 44.0%; Pred. No. 3.1e-22;

Matches 95; Conservative 39; Mismatches 61; Indels 21; Gaps 7;
QY 1 IKVNMWDLFFSPSEDFNTDNLKNGEITSDPTNEAENISLDLQGYLTFNFDNPE 59
DB 436 IKVNMWDLFFVASSESYNENGINSPKEI-DPTTINNNYKKMLD---EVLIDYNSD---- 487

ID	Accession	Species	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	60	NIENISNLSDDIGOL----	ELMPNTEREPNG----	KKYLDTYTFHRYRAOEFHGKSR	111	111	111	111	111	111
Db	488	--AIPNLSRLLINTAAQNDSDYVPKID--	SNGISEIKETIVDLINVFYIYAAKAPBGESA	543	543	543	543	543	543	543
QY	112	IALTNSVNEALLNPSRVYTFPSSDYVKVKNKATFAMFLGWEOLVYDFTDSEVSTD	171	171	171	171	171	171	171	171
Db	544	ISLTSSVNTALLDASKAYTFEFSDDPINFVKNRPVQAALFISWIOQYINDETBTATQKSTD	603	603	603	603	603	603	603	603
QY	172	KIADITIIIPYIGPALNTGNMLYRKDFGALIFSGA	207	207	207	207	207	207	207	207
Db	604	KIADISLIVPYIGLALNTGNENQKNFKAELIELLGA	639	639	639	639	639	639	639	639

RESULT 2

Q9FAR6 PRELIMINARY: PRT: 1255 AA.

Q9FAR6 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

GN Type E botulinum toxin.

GN BONT/E.

OS Clostridium butyricum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;

OC Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxID:1492;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BL 6340/ATCC 43755/BL 5520/KZ 147;

RX MEDLINE=20509829; PubMed=11055954;

RA Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y.,

RA Yanakawa K., Oguma K., Sakaguchi Y., Nakamura S.;

RT "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium

RT butyricum Strains."

RL Appl. Environ. Microbiol. 66:4992-4997(2000).

DR EMBL: AB039264; BABI2249.1; -.

DR HSSP; P10845; 3BTA.

DR InterPro: IPR000395; Bontoxilysin.

DR InterPro: IPR000130; Zn_MRPepidse.

DR Pfam: PF01742; Peptidase_M27; 1.

DR PRINTS: PR00760; BONTOXILYSIN.

DR ProDom: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.

SQ SEQUENCE 1255 AA; 143918 MW; 1B557B9D85CD8E4D CRC64;

Query Match 35.0%; Score 374.5; DB 2; Length 1255;

Best Local Similarity 38.9%; Pred. No. 9.2e-20;

Matches 84; Conservative 51; Mismatches 60; Indels 21; Gaps 7.

ID	Accession	Species	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
QY	1	IKVNMWDLFFSSSEDNFTND-LNKGEI----	TSDTNIEMAEENISLIDLIQYVTFNFD	55	55	55	55	55	55	55
Db	430	IEINNGELFEVASENSYNDNINFPKEDIPTYSNNNE-----	NLDQVIINPNE	481	481	481	481	481	481	481
QY	56	NEPENISIEENLSSDITIGOLELMPNIERRPPNG----	KKYLDTYTFHRYRAOEFHGKSR	111	111	111	111	111	111	111
Db	482	SAP-GLSDEKLLTLTQND-AVLPKYD--	SNGTSIDIEOHVNEELNVEFYLDAAKQVDEGENN	537	537	537	537	537	537	537
QY	112	IALTNSVNEALLNPSRVYTFPSSDYVKVKNKATFAMFLGWEOLVYDFTDSEVSTD	171	171	171	171	171	171	171	171
Db	538	VNLTSIDIDALLEQPKITTFEFSSEFINNNKVPQAALFGWIOQYINDETBTATQKSTD	597	597	597	597	597	597	597	597
QY	172	KIADITIIIPYIGPALNTGNMLYRKDFGALIFSGA	207	207	207	207	207	207	207	207
Db	598	KIADISIVPYIGLALNTGNENQKNFKAELIELLGA	633	633	633	633	633	633	633	633

RESULT 3

Q9K395 PRELIMINARY: PRT: 1251 AA.

Q9K395 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
DT Type E botulinum toxin.
GN BONT/E.
GN Clostridium butyricum.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 095;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (LCL 095) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 155;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K
RT "C. butyricum (LCL 155) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1899;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1899) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1897;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1897) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1898;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1898) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1886;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1886) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1887;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1887) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1889;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1889) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1890;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RT "C. butyricum (KZ 1890) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1891;

```

RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Katasawa T.,
RT "C. butylicum (K2 1891) gene for type E botulinum toxin.",
RT RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=LCI 063;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Katasawa T.,
RT "C. butylicum (LCI 063) gene for type E botulinum toxin.",
RT RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB037714; BAB03522.1; -
DR EMBL: AB037704; BAB03512.1; -
DR EMBL: AB037705; BAB03513.1; -
DR EMBL: AB037706; BAB03514.1; -
DR EMBL: AB037707; BAB03515.1; -
DR EMBL: AB037708; BAB03516.1; -
DR EMBL: AB037709; BAB03517.1; -
DR EMBL: AB037710; BAB03518.1; -
DR EMBL: AB037711; BAB03519.1; -
DR EMBL: AB037712; BAB03520.1; -
DR EMBL: AB037713; BAB03521.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -
DR InterPro; IPR000395; Bontotoxylisin.
DR InterPro; IPR000130; Zn_Metpeptase.
DR Pfam; PF01742; Peptidase_M27.1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR PRODOM; PD001963; Bontotoxylisin; 1.
DR PROSITE; PS00142; ZINC_PROTASE; UNKNOWN_1.
DR SEQUENCE 1251 AA; 143751 MW; 2021F4E427070296 CRC64;

```

Query Match	34.44:	Score 368.5:	DB 2:	Length 1251:
Best Local Similarity	38.44:	Pred. NO.2.0e-19:		
Matches	83:	Conservative 51:	Mismatches 61:	Indels 21:
Gaps				
Qy	1	IKVNMWDFEFSPESDNFTND-LANGEEI---	TSDTIEAAEENISDLIQOYYLTPEFD	55
Dy	427	IEINNGLFEFYAASENSYNDNINPKELDQVTSNNNYE-----	NDDQVILNNNSE	478
Qy	56	NEPERNISLENISDLIGOLELMPNIEERPNG---	KKYEIDKRYTMFYHRAQEFHGKSR	111
Dy	479	SAP-GLSDEKLNLTIOND-AVIPKPYD--	SNQTSIDIEQHDVNELVFEFFYLDAQKPEGENN	534
Qy	112	IALNVSVEALNINSRYATPEFSSSYVKKNKATFAAFLGMEVLYVDPFDEISEVSTTD	171	
Dy	535	VNLTSIDTALLEOPKTYTFEFSSEFINNVKNPVQOAAFLVSMIOQVLDVFTTEANOKSTVD	594	
Qy	172	KIADTITIPYIGPALNIGNNLKYDDPFGVALIFSGA	207	
Dy	595	KIADISIVPYIGTALNIGNNAQGNKRKDALELLGA	630	
RESULT 4				
Q933KO		PRELIMINARY;	PRT; 1291 AA.	
Q933KO				
AC	0933KO:			
DT	01-DEC-2001 (TReMBLrel. 19, Created)			
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)			
DE	TYPE B cryptic neurotoxin.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;			
OC	Clostridiales; Clostridiaceae; Clostridium.			
OX	NCB1_TaxID=1491;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	SPRAIN=593, AND 588;			
RC	Kirma N., Ferreira J.L., Baumstark B.R.;			
RA	"Characterization of six type A strains of Clostridium botulinum that			
RT	contain type B toxin gene sequences.";			
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.			
RR	EMBL: AF300466; AAL1499.1; -			

DR EMBL: AF300465; AALL1498.1; -
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTEPpepde.
DR Pfam: PF01742; Peptidase_M27; 1.
DR Prodom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin.
SQ SEQUENCE 1291 AA; 150843 MW; 7AC1737B0FA5A151 CRC64;

Query Match	34.3%	Score 367.5;	DB 2;	Length 1291;
Best Local Similarity	38.3%;	Pred. No. 3.2e-19;		
Matches	85;	Conservative	40;	Mismatches 62; Indels 35; Gaps 3;
QY	1	IKVNWMDLEFFSSSEDNFENDLNKGEETISDPNIEAAEENISL-DLIQOYLYTFNEDNEPE	59	
		: : : : : : : : : :		
Db	447	IDVDNEDLEFFADKNKSFSDLLSKNERIAYNQNNTIENDFSINELI-----	492	
QY	60	NISIMLSSDIIGQEL-----MPNIERPPNGKKYILDKYTMFHYRAQEF	105	
		: : : : : : : : : : : : : : : : : : : :		
Db	493	-----LDTDLISKLELSESENTESLTJDENYVYVYKKQPAIKLFTEDNTLEQYLYSQTF	546	
QY	106	EHGKSRILATSVNVAENLNPSSRVYTFSSDYVKKYNAKTAAMLEAGWQOYLVEFTDTS	165	
		: : : : : : : : : : : : : : : : : : : :		
Db	547	PLIDIDISLTFSDALLFNKMYSFSEMDIKTKANKKYVEAGLFRAGWKQYLNDVFIEAN	606	
QY	166	EVSYTDKADITITIIIPYIGPALINIGNMLYKODENAGALISQA	207	
		: : : : : : : : : : : : : : :		
Db	607	KSSYMDKTAIDISLIVPYIGLALNGENETAKGFEENAFIAGA	648	

ID	093G71	PRELIMINARY;	PRT; 1291 AA.
AC	093G71.		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, last annotation update)		
DE	Neurotoxin type B.		
OS	Clostridium botulinum.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;		
OC	Clostridiales; Clostridiaceae; Clostridium.		
OX	NCBI_TaxID=1491;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=1436;		
RA	Kirma N., Ferreira J.L., Baumstark B.R.;		
RT	"Characterization of six type A strains of Clostridium botulinum that		
RT	contain type B toxin gene sequences.";		
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF295926; AAK97132.1; -		
DR	InterPro: IPR000395; Bontoxilysin.		
DR	InterPro: IPR000130; Zn_MpPepIdc.		
DR	Pfam; PF01742; Peptidase_M27.1.		
DR	ProDom; PD001965; Bontoxilysin; 1.		
DR	PROSITE; PS00142; ZINC-PROTASE; UNKNOWN_1.		
QO	SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EB8CD2 CRC64;		

	Query Match	34.2%	Score	366.5;	DB 2:	Length	1291;			
	Best Local Similarity	38.5%;	Pred.	No. 3.8e-19;						
	Matches	85;	Conservative	41;	Mismatches	62;	Indels	33;	Gaps	3;
Qy	1	IKVNWDLFFSPSENFINDLNKGEETSDTIEAAEENISLDLIQQYYLFPNPONEPEN	60	: : : : : :						
Dd	447	IDVDNEDLFLLADKKSFSDLSKNERIAVT-----QNNTIDNF-----S	487	: : : : : :						
Qy	61	ISIENI.SDIIIGOLEL-----MENIERFPNGKKYELDKYTMEHYLRAGEFE	106	: : : : : :						
Dd	488	INELLTDPLDTLSKIELPSENTESLDFDNVVYPPEYKKOFAIKKIIPFDEMTIFQYLVSQTFP	547	: : : : : :						
Qy	107	HGSKSIALTNSVENMLMPSKRYITFTSSDYAKKKVKATAEAMFLCWEGOLVYDFPDEISE	166	: : : : : :						
Dd	548	LDIRISLTSTSSDDALLSNKYISFFESNDYIKTKARVVEAGFLFGAQVQIYDDFVEANK	607	: : : : : :						

[illegible]

RA Kimura B., Kawasaki S., Nakano H., Fujii T.;
 RT "Type B in Modified-Atmosphere Packaging of Growth of Clostridium botulinum
 RT Appl. Environ. Microbiol. 67:206-216(2001)."
 DR EMBL: AB040126; BAB07888.2; -;
 DR HSSP: P10845; 3BTA.
 KW Neurotoxin.
 FT NON_TER 1 1
 FT NON_TER 173 173
 SQ SEQUENCE 173 AA; 19137 MW; 4178024F5BC4BF1F CRC64;

Query Match 30.8%; Score 330; DB 2; Length 173;
 Best Local Similarity 35.7%; Pred. No. 1.8e-17;
 Matches 71; Conservative 48; Mismatches 48; Indels 32; Gaps 5;

OY 13 SDNFTNDLNKGEITSQNIIEA---EENISLDLQOYLTFFNDNEPENISTENSS 68
 DB 1 SNNNNENDL---DQVTLNFSSESAPGLSDEKLNLTQNDAYIP-KYDS-----NGTS 48
 OY 69 DIIGOLELMPNIEFPGNGKYEIDKTYTMFHYLRAOEFEGKSRIALTNSVNEALNPSRY 128
 DB 49 DI-----BOHDVNEINVFYIDAKVPEGENNVNLTSSIDTRALLLPQKI 92
 OY 129 YTFESSDYVKVKNKATEAMFLGWEOQVYDFTDESEVSTTDKIADITIIPIYIGPALN 188
 DB 93 YTFESSSEFINNVKRPQALFVGWIOQVLDVFTTEANQESTVDKIDISIVPYKGLALN 152
 OY 189 IGNNMLYKDDFVAGALIFSGA 207
 DB 153 IGNEAKGNGKDALLELLGA 171

RESULT 14

O93N27 PRELIMINARY; PRT; 1310 AA.
 AC O93N27;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Tetanus toxin (Fragment).
 OS Clostridium tetani.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shumin Z., Dianliang L.;
 RT "Cloning and sequence analysis of tetanus toxin gene."
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF389424; AAK72964.2; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR000130; Zn_MpPeptidase.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 FT NON_TER 1 1
 FT NON_TER 1310 1310
 SQ SEQUENCE 1310 AA; 150316 MW; 9EADDC914418EA50 CRC64;

Query Match 29.3%; Score 314; DB 2; Length 1310;
 Best Local Similarity 33.2%; Pred. No. 3.4e-15;
 Matches 74; Conservative 44; Mismatches 67; Indels 38; Gaps 4;

OY 1 IKVNMNDLFSSSEDNFTNDLNKGEITSQNIIEAENISLDLQOYLTFFNFDNEPEN 60
 DB 469 IKIKNEDLFIKKNSESEPPODEIVSYNTKKNPLNFYSLDKITLDY----- 517
 OY 61 ISTEINSSDIIGOLELMPNIEFPGNGK-----YELDKYTMFHYLRAQ 103
 DB 518 ----NLQSKI-----TLPNDRTTPYTKGIPYAPKESNAASTLEIHNDIDNTIYQYLYAQ 568

OY 104 EEFHGKSRIALTNSVNEALNPSRYTFFSDYVKVKNKATEAMFLGWEOQVLDFTDE 163
 DB 569 KSTTITQRTITMINSVDALLNKIKISYFPP-VISKVNGAGQILFLQVNRDIDFTNE 627
 OY 164 TSEVSTTDKIADITIIPIYIGPALNIGNNMLYKDDFVAGALIFSG 206
 DB 628 SSQKTTIDKISDVSTIVPYIGPALNIVKOGYEGNFGALFTTG 670

RESULT 15

O9LBR1 PRELIMINARY; PRT; 1285 AA.
 AC O9LBR1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Neurotoxin.
 GN NT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D-4947;
 RA Sagane Y., Watanabe T., Kouguchi H., Yamamoto T., Takizawa J.,
 RA Kawabe T., Murakami F., Muroga A., Nakatsuka M., Ohyama T.;
 RT "Characterization of the Progenitor Toxin Components Produced by
 RT Clostridium botulinum Type D Strain 4947."
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB037920; BAA90661.1; -;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MpPeptidase.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1285 AA; 147352 MW; B63AFA487D570680 CRC64;

Query Match 28.3%; Score 303.5; DB 2; Length 1285;
 Best Local Similarity 33.2%; Pred. No. 2e-14;
 Matches 72; Conservative 38; Mismatches 80; Indels 27; Gaps 4;

OY 1 IKVNMNDLFSSSEDNFTNDLNKGEITSQNIIEAENISLDLQOYLTFFNFDNEPEN 60
 DB 451 IQVKNNTLPLYVADKDISOEIESQITDETENVENSDFNSLD---ESTIDAKVPTNPEA 507
 OY 61 ISTEINSSDIIGOLELMPNI-----EFPFGKKYEIDKTYTMFHYLRAOEFHGK 109
 DB 508 VD-----PLLPVNMNEPLNVPGEEVEFYDITKVDVYLNSTYLLKOKLSNV 555
 OY 110 SRIALTNVNEALNPSRYTFFSDYVKVKNKATEAMFLGWEOQVLDVFTDESEVST 169
 DB 556 ENTTLTSVEALGYNKTKYTFPLS-LAKVKNVGAGALFLMNAANVDEFTTNINKKOT 614
 OY 170 TDKIADITIIPIYIGPALNIGNNMLYKDDFVAGALIFSG 206
 DB 615 LDKISDVSAIIPYIGPALNIGNSALGNFKQAFATAG 651

Search completed: March 13, 2003, 11:40:11
 Job time : 35.5202 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:12 ; Search time 26.2248 Seconds
(without alignments) 711.353 Million cell updates/sec

Title: US-09-917-791-22

Perfect score: 727

Sequence: 1 LNSSLVKGTFIIKKYASGN.....QIERSRTLCGSMFIPVD 140

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	100.0	140	22	AAG79296
2	727	100.0	206	21	AA77144
3	727	100.0	382	21	AA836303
4	727	100.0	415	22	AA804083
5	727	100.0	432	21	AA77142
6	727	100.0	434	22	AA804089
7	727	100.0	435	22	AA804090
8	727	100.0	437	22	AA804088
9	727	100.0	438	17	AA85008
10	727	100.0	438	19	AA868389

11	727	100.0	438	21	AA77134
12	727	100.0	445	19	AA868391
13	727	100.0	462	17	AA89509
14	727	100.0	462	19	AA868390
15	727	100.0	837	21	AA77140
16	727	100.0	1067	21	AA93307
17	727	100.0	1092	21	AA93310
18	727	100.0	1296	17	AA895010
19	701	96.4	1295	23	AA895339
20	698.5	96.1	847	22	AA804081
21	284.5	39.1	432	21	AA77138
22	277.5	38.2	143	18	AA809017
23	277.5	38.2	431	18	AA809014
24	277.5	38.2	432	22	AA804093
25	277.5	38.2	432	22	AA804105
26	277.5	38.2	645	22	AA807894
27	277.5	38.2	685	22	AA807893
28	277.5	38.2	862	22	AA807890
29	277.5	38.2	887	22	AA807892
30	277.5	38.2	1032	22	AA807901
31	277.5	38.2	1059	21	AA93309
32	277.5	38.2	1084	21	AA93312
33	277.5	38.2	1092	22	AA807900
34	258.5	35.6	448	19	AA868399
35	206.5	28.4	419	22	AA804095
36	206.5	28.4	449	21	AA77137
37	206.5	28.4	449	22	AA804094
38	206.5	28.4	452	19	AA868396
39	190	26.1	451	19	AA868395
40	169.5	23.3	440	21	AA77135
41	169.5	23.3	440	22	AA804091
42	169.5	23.3	472	19	AA868393
43	165.5	22.8	472	19	AA868394
44	165.5	22.8	1070	21	AA93308
45	165.5	22.8	1095	21	AA93311

ALIGNMENTS

RESULT 1
AAG79296 standard; Protein: 140 AA.

AC AAG79296;
03-JAN-2002 (first entry)

DE Amino acid sequence of botulinum neurotoxin (BoNT) fragment.
KW Neurotoxin; BoNT; vaccine; botulism.
OS Clostridium botulinum.
PN US6287566-B1.
PD 11-SEP-2001.
PF 19-MAY-1995; 95US-0446114.
PR 19-MAY-1995; 95US-0446114.
PA (USSA) US SEC OF ARMY.
PI Dertzbauhg MT.
PS WPI: 2001-615462/71.
New protective epitopes of neurotoxin of Clostridium botulinum, useful for developing vaccines against neurotoxins of Clostridium botulinum and for identifying protective antibodies -
Claim 1; Column 25; 14pp; English.

of picard

Synthetic botulinum
Clostridium botuli
Type A neurotoxin
Clostridium botuli
Native botulinum n
A manganeese supero
C. botulinum type
Clostridium botuli
Botulinum toxin hea
Synthetic botulinu
Immunogenic type F
Botulinum toxin hea
Botulinum toxin hea
Modified clostridi
Modified clostridi
Modified clostridi
Modified clostridi
C. botulinum C2 tr
A manganeese supero
A manganeese supero
C. botulinum C2 tr
Clostridium botuli
Clostridium botuli
Synthetic botulinu
Botulinum toxin hea
Clostridium botuli
Clostridium botuli
A manganeese supero

XX The present sequence represents a fragment of the Clostridium botulinum
CC neurotoxin (BoNT). It was produced by amplifying overlapping fragments
CC of the BoNT gene. The amplified fragments were cloned expressed to
CC identify immunogenic polypeptides which are capable of giving rise to
CC protective antibodies. The BoNT polypeptide fragment are useful as
CC vaccines, for immunizing against botulism, and as diagnostic agents
CC to identify protective antibodies.
XX
SQ Sequence 140 AA:
Query Match 100.0%; Score 727; DB 22; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LNSSLYRGTFIKKYASGKNDIVRNDRYINVVVKKEYRLATNASQAGEKILSL 60
DB 1 LNSSLYRGTFIKKYASGKNDIVRNDRYINVVVKKEYRLATNASQAGEKILSL 60
QY 61 EIPDVGNLSQVYVVKSKNDGITNKCKMNLQDNGNDIGFIGHOFNNIAKLVA5MWYNR 120
DB 61 EIPDVGNLSQVYVVKSKNDGITNKCKMNLQDNGNDIGFIGHOFNNIAKLVA5MWYNR 120
QY 121 QIERSSRTLCGSWEFTIPVD 140
DB 121 QIERSSRTLCGSWEFTIPVD 140
RESULT 2
AAV77144
ID AAV77144 standard; Protein; 206 AA.
XX
AC AAV77144;
XX
DT 08-MAY-2000 (first entry)
XX
DE Botulinum neurotoxin serotype A (BoNTA) C-terminal subfragment AsubHc2.
XX
KW Botulinum neurotoxin; heavy chain; BoNT; serotype A;
KW C-terminal subfragment; Hc; botulism; VEE;
KW Venezuelan equine encephalitis virus replicon; vaccine; diagnosis;
KW drug screening.
XX
OS Clostridium botulinum.
OS Synthetic.
XX
PN WO200002524-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US15570.
XX
PR 10-JUL-1998; 98US-0092416.
PR 12-MAY-1999; 99US-0133870.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX
DR MPI: 2000-160827/14.
DR N-PSDB; AA287222.
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
PS Disclosure; Page 53; 54pp; English.
XX
CC The invention relates to novel vaccines that induce a protective immune
CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT

CC serotypes A-G. In preferred embodiments of the invention, the vector is a
CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC vector results in the production of large amounts of a protein encoded by
CC a sequence cloned into the replicon. The constructs are used to produce
CC vaccine against botulism. The proteins can also be used as diagnostic
CC tools for the diagnosis of botulism. The transformed host cells can be
CC used to analyse the effectiveness of drugs and agents which inhibit toxin
CC effects. The vaccine currently used against botulism is dangerous
CC and expensive to produce, and contains formalin, which is very painful
CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC the 7 serotypes are represented in the formulation. The novel vaccine
CC of overcomes these problems, as it is easily purified, and available in
CC large quantities. It is also expressed in the lymph nodes for a better
CC immune response. The present sequence represents BoNTA heavy chain
CC C-terminal subfragment AsubHc2, comprising residues 234 to 438 of the
CC BoNTA Hc fragment plus an initial methionine, and was used in the present
CC invention.
XX
SQ Sequence 206 AA:
Query Match 100.0%; Score 727; DB 21; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.2e-77;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LNSSLYRGTFIKKYASGKNDIVRNDRYINVVVKKEYRLATNASQAGEKILSL 60
DB 60 LNSSLYRGTFIKKYASGKNDIVRNDRYINVVVKKEYRLATNASQAGEKILSL 119
QY 61 EIPDVGNLSQVYVVKSKNDGITNKCKMNLQDNGNDIGFIGHOFNNIAKLVA5MWYNR 120
DB 120 EIPDVGNLSQVYVVKSKNDGITNKCKMNLQDNGNDIGFIGHOFNNIAKLVA5MWYNR 179
QY 121 QIERSSRTLCGSWEFTIPVD 140
DB 180 QIERSSRTLCGSWEFTIPVD 199
RESULT 3
AAB36303
ID AAB36303 standard; Protein; 382 AA.
XX
AC AAB36303;
XX
DT 15-FEB-2001 (first entry)
XX
DE BoNT/A prototoxin heavy chain C-terminal neural cell binding domain.
XX
KW Human; procholecystokinin; CCK A receptor; CCK B receptor;
KW pancreatitis; antiinflammatory.
XX
OS Clostridium botulinum.
XX
PN WO200061192-A2.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09142.
XX
PR 08-APR-1999; 99US-0288326.
XX
PA (ALIR) ALLERGAN SALES INC.
XX
PI Steward LE, Sachs G, Aoki KR;
XX
DR MPI: 2000-679416/66.
XX
PT New composition for treating acute pancreatitis, comprises a pancreatic
PT cell surface marker binding element, a translocation element that
PT transfers polypeptide across vesicular membrane, and a therapeutic
PT element -
XX
PS Example 1; Page 29; 50pp; English.
XX

CC The present invention describes a composition (I) for treating acute
 CC pancreatitis. (I) comprises a first element containing a binding element
 CC that binds to a pancreatic cell surface marker, a second element
 CC containing a translocation element that facilitates polypeptide transfer
 CC across the vesicular membrane, and a third element containing a
 CC therapeutic element that inhibits enzyme secretion in pancreatic cell
 CC cytoplasm. Also described is a method for making a therapeutic
 CC polypeptide having a binding element selective for cholecystokinin (CCK)
 CC receptor by expressing within a host cell a recombinant chimeric
 CC polypeptide comprising an extein containing a therapeutic element and a
 CC translocational element, and an intein located to the carboxy terminal
 CC of extein having a cysteine, serine or threonine at its amino terminus,
 CC and contacting the extein with a synthetic peptide comprising a CCK
 CC amino acid sequence containing an amidated phenylalanine at a natural
 CC C-terminus, and a cysteine, serine or threonine at its N-terminus, and
 CC a nucleophilic reagent able to cause cleavage of the intein to form a
 CC peptide bond between the extein C-terminus and synthetic peptide
 CC N-terminus through the formation of an activated ester or thio ester
 CC intermediate. (I) has antiinflammatory activity and prevents accumulation
 CC of pancreatic digestive enzymes, and prevents exocytic fusion of vesicles
 CC containing secretory enzymes of pancreas. (I) is useful for treating
 CC acute pancreatitis. The present sequence represents the Clostridium
 CC botulinum BoNT/A prototoxin heavy chain C-terminal neural cell binding
 CC domain, which is given in the exemplification of the present invention.
 CC
 XX SQ Sequence 382 AA:

Query Match 100.0%; Score 727; DB 21; Length 382;
 Best Local Similarity 100.0%; Pred. No. 5.6e-77;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNSSLYRGKFKFIKKKASGKNDIVNRNRYTINVVYKKEKRLATNASOAGYEKILSL 60
 DB 236 LNSSLYRGKFKFIKKKASGKNDIVNRNRYTINVVYKKEKRLATNASOAGYEKILSL 295
 OY 61 EIPDVGNLSQVYVYVMSKNDGITNCKKMNLDNNGNDIGFIGHQPNNTAKLIVASWMYNR 120
 DB 296 EIPDVGNLSQVYVYVMSKNDGITNCKKMNLDNNGNDIGFIGHQPNNTAKLIVASWMYNR 355
 OY 121 QIERSSRTLGCSWEFIPVD 140
 DB 356 QIERSSRTLGCSWEFIPVD 375

RESULT 4

AAB04083 standard; protein: 415 AA.

AC AAB04083;
 XX
 DT 11-APR-2001 (first entry)
 XX
 DE Botulinum toxin C fragment sequence (serotype A).
 XX
 KW Botulinism; toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KW infection.
 XX
 OS Clostridium botulinum.
 XX
 PN WO200067700-A2.
 PD 16-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US12890.
 XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.

XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PA Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 PI WPI: 2001-016048/02.
 DR
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulinism
 XX
 PS Example 7; Page 36; 73pp; English.

CC Botulin neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.

SQ Sequence 415 AA:

Query Match 100.0%; Score 727; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 6.3e-77;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNSSLYRGKFKFIKKKASGKNDIVNRNRYTINVVYKKEKRLATNASOAGYEKILSL 60
 DB 269 LNSSLYRGKFKFIKKKASGKNDIVNRNRYTINVVYKKEKRLATNASOAGYEKILSL 328
 OY 61 EIPDVGNLSQVYVYVMSKNDGITNCKKMNLDNNGNDIGFIGHQPNNTAKLIVASWMYNR 120
 DB 329 EIPDVGNLSQVYVYVMSKNDGITNCKKMNLDNNGNDIGFIGHQPNNTAKLIVASWMYNR 388
 OY 121 QIERSSRTLGCSWEFIPVD 140
 DB 389 QIERSSRTLGCSWEFIPVD 408

RESULT 5

AAV77142 standard; protein: 432 AA.

AC AAV77142;
 XX
 DT 08-MAY-2000 (first entry)
 XX
 DE Native botulinum neurotoxin serotype A (BoNTA) C-terminal fragment (HC).
 XX
 KW Botulinum neurotoxin; heavy chain; BoNT; serotype A;
 KW C-terminal fragment; Hc; Venezuelan equine encephalitis virus replicon;
 KW VEE; botulinism; vaccine; diagnosis; drug screening.
 XX
 OS Clostridium botulinum.
 XX
 PN WO200002524-A2.
 PD 20-JAN-2000.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 432 /note="Apparently encoded by GGATGGCGAG AAGGCCACT G"
 XX
 PN
 PD
 XX

```

PE 09-JUL-1999;          99WO-US15570.
PR 10-JUL-1998;         98US-0092416.
PR 12-MAY-1999;         99US-0133870.
XX                                     (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
PA
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
DR WPI: 2000-160827/14.
DR N-PSDB; AAZ87220.
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
PS Example 3; Page 52; 54pp; English.
XX
CC The invention relates to novel vaccines that induce a protective immune
CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT
CC serotypes A-G. In preferred embodiments of the invention, the vector is a
CC Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this
CC vector results in the production of large amounts of a protein encoded by
CC a sequence cloned into the replicon. The constructs are used to produce
CC vaccines against botulinism. The proteins can also be used as diagnostic
CC tools for the diagnosis of botulinism. The transformed host cells can be
CC used to analyse the effectiveness of drugs and agents which inhibit toxin
CC effects. The vaccine currently used against botulinism is dangerous
CC and expensive to produce, and contains formalin, which is very painful
CC for the recipient. Also, the vaccine is incomplete, in that only 5 of
CC the 7 serotypes are represented in the formulation. The novel vaccine
CC of overcomes these problems, as it is easily purified, and available in
CC large quantities. It is also expressed in the lymph nodes for a better
CC immune response. The present sequence represents the native BoNTA heavy
CC chain C-terminal fragment (Hc) used in an exemplification of the present
CC invention.
CC
XX Sequence 432 AA;
XQ
Query Match      100.0%; Score 727; DB 21; Length 432;
Best Local Similarity 100.0%; Pred. No. 6,7e-77;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 1 LNSSLYRGKFLITKKKYASGNKDINVRNDRYYINVVVKKREYLATNASAGYEKLISAL 60
DB 232 LNSSLYRGKFLITKKKYASNKDINVRNDRYYINVVVKKREYLATNASAGYEKLISAL 321
OY 61 ELIPDVGNLSQVVVMVMSKNQGITNKCKMMLQDNNGNDIGFIGFHOPNNIAKLVAASWMYNR 120
DB 352 ELIPDVGNLSQVVVMVMSKNQGITNKCKMMLQDNNGNDIGFIGFHQNNTIAKLVAASWMYNR 411
OY 121 QIERSSRTLGCSWEFIYPVD 140
DB 412 QIERSSRTLGCSWEFIYPVD 431
RESULT 6
AAB04089
AACB04089 standard; Protein; 434 AA.
AACB04089;
DT 11-APR-2001 (first entry)
DE Botulinism toxin heavy chain C-terminal sequence (serotype A).
KW Botulinism; toxin; neurotoxin; heavy chain; recombinant expression;
recombinant vector; antigen; immune response; vaccine; bacterium;
infection.

```

Query Match	100.0%;	Score 727;	DB 22;	Length 434;
Best Local Similarity	100.0%;	Pred. No. 6,7e-77;		
Matches 140;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1 LNSSLYRGKFIKKIYASGKNKDIYRNDRIYINVVYKKERLATMNSQAQVERKLSL 60				
DB 288 LNSSLYRGKFIKKIYASGKNKDIYRNDRIYINVVYKKERLATMNSQAQVERKLSL 347				
QY 61 ETIPDGNLSQVYVMSKNDGITTNCCKMNLQNNNGNDIGFGEFHQFNNTAKLVASWNYR 120				
DB 348 ETIPDGNLSQVYVMSKNDGITTNCCKMNLQNNNGNDIGFGEFHQFNNTAKLVASWNYR 407				
QY 121 QIERSSRTLGCSEWFEIPVD 140				
DB 408 QIERSSRTLGCSEWFEIPVD 427				
RESULT 7				
AAB04090				
ID AAB04090 standard; Protein; 435 AA.				
AC AAB04090;				
XX				
XX				
DT 11-APR-2001 (first entry)				

```

XX DE Botulism toxin heavy chain C-terminal sequence (serotype A).
XX XX
XX Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
XX KM recombinant vector; antigen; immune response; vaccine; bacterium;
XX KM infection.
XX OS Synthetic.
XX OS Clostridium botulinum.
XX PN WO200067700-A2.
XX PD 16-NOV-2000.
XX PF 12-MAY-2000; 2000WO-US12890.
XX PR 12-MAY-1999; 99US-0133865.
XX PR 12-MAY-1999; 99US-0133866.
XX PR 12-MAY-1999; 99US-0133867.
XX PR 12-MAY-1999; 99US-0133868.
XX PR 12-MAY-1999; 99US-0133869.
XX PR 12-MAY-1999; 99US-0133870.
XX PR 29-JUL-1999; 99US-0146192.
XX PA (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX DR WPI: 2001-016048/02.
XX DR N-PSDB; AAA54484.
XX XX
XX PT New nucleic acids encoding the carboxy- or amino-terminal portions of
XX PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
XX PT vaccine against botulism
XX PS Disclosure: Fig 3b; 73pp; English.
XX PS
XX CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
XX CC chain and then posttranslationally nicked, forming a dichain
XX CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
XX CC remain linked by a disulfide bond. Nucleic acids encoding the
XX CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
XX CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
XX CC expression vectors and expressed in transformed cells to produce
XX CC peptide antigens useful for eliciting an immune response to give
XX CC protective immunity against botulinum neurotoxin, which causes
XX CC botulism. The nucleic acids are expressible in a recombinant
XX CC organisms such as Escherichia coli or Pichia pastoris. The use
XX CC of recombinant nucleic acids are advantageous since it eliminates
XX CC the need to culture large quantities of hazardous toxin-producing
XX CC bacterium. Production yield from the genetically engineered product
XX CC is also high and cost of production is lower. The nucleic acids can
XX CC be derived from Clostridium botulinum serotypes A-G.
XX XX
XX SQ Sequence 435 AA;
XX
XX Query Match 100.0%; Score 727; DB 22; Length 435;
XX Best Local Similarity 100.0%; Pred. No. 6.8e-77;
XX Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESUL 8
AAB04088
XX ID AAB04088 standard; Protein; 437 AA.
XX AC AAB04088;
XX XX
XX DT 11-APR-2001 (first entry)
XX DE Botulism toxin heavy chain C-terminal sequence (serotype A).
XX KM Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
XX KM recombinant vector; antigen; immune response; vaccine; bacterium;
XX KM infection.
XX OS Synthetic.
XX OS Clostridium botulinum.
XX PN WO200067700-A2.
XX PD 16-NOV-2000.
XX PF 12-MAY-2000; 2000WO-US12890.
XX PR 12-MAY-1999; 99US-0133865.
XX PR 12-MAY-1999; 99US-0133866.
XX PR 12-MAY-1999; 99US-0133867.
XX PR 12-MAY-1999; 99US-0133868.
XX PR 12-MAY-1999; 99US-0133869.
XX PR 12-MAY-1999; 99US-0133870.
XX PR 29-JUL-1999; 99US-0146192.
XX XX
XX PA (USSA ) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX DR WPI: 2001-016048/02.
XX DR N-PSDB; AAA54482.
XX XX
XX PT New nucleic acids encoding the carboxy- or amino-terminal portions of
XX PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
XX PT vaccine against botulism
XX PS Claim 3; Fig 1b; 73pp; English.
XX PS
XX CC Botulism neurotoxins are translated as a single 150 kDa polypeptide
XX CC chain and then posttranslationally nicked, forming a dichain
XX CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
XX CC remain linked by a disulfide bond. Nucleic acids encoding the
XX CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
XX CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
XX CC expression vectors and expressed in transformed cells to produce
XX CC peptide antigens useful for eliciting an immune response to give
XX CC protective immunity against botulinum neurotoxin, which causes
XX CC botulism. The nucleic acids are expressible in a recombinant
XX CC organisms such as Escherichia coli or Pichia pastoris. The use
XX CC of recombinant nucleic acids are advantageous since it eliminates
XX CC the need to culture large quantities of hazardous toxin-producing
XX CC bacterium. Production yield from the genetically engineered product
XX CC is also high and cost of production is lower. The nucleic acids can
XX CC be derived from Clostridium botulinum serotypes A-G.
XX XX
XX SQ Sequence 437 AA;
XX
XX Query Match 100.0%; Score 727; DB 22; Length 437;
XX Best Local Similarity 100.0%; Pred. No. 6.8e-77;
XX Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

DB 351 EIPDVGNLSQVYVVKSKNDGINTKCKMNLQDNNNGNDIGFIGHQFNNTAKLVA\$MWYNR 410
 QY 121 QIERSSRTLGC\$SWEFI\$PVD 140
 DB 411 QIERSSRTLGC\$SWEFI\$PVD 430

RESULT 9
 AAR95008
 ID AAR95008 standard; Protein; 438 AA.

XX AAR95008;
 XX
 DT 07-JUL-1996 (first entry)
 XX
 DE Type A neurotoxin C fragment.

XX Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;
 KW Clostridium botulinum.

OS Synthetic.

PN WO9612802-A1.

PD 02-MAY-1996.

PF 23-OCT-1995; 95WO-US13737.

PR 07-JUN-1995; 95US-0480604.

PR 24-OCT-1994; 94US-0329154.

PR 16-MAR-1995; 95US-0405496.

PR 14-APR-1995; 95US-0422711.

XX (OPHI-) OPHIDIAN PHARM INC.

PI Pirca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;

PI Williams JA;

DR WPI: 1996-230603/23.

DR N-PSDB; AAT929245.

XX Fusion proteins comprising non-toxin protein and part of toxin

PT useful to form anti-toxins against Clostridium botulinum type A, and

PT C. difficile type toxins, and to treat C. difficile intoxication,

PT paritc. diarrhoea

XX

PS Claim 5; Page 336-38; 434pp; English.

XX The heavy chain C fragment (AAR95008) of Clostridium botulinum type

CC A neurotoxin (see also AAR95010) was produced by expression of a

CC synthetic gene (AAT9245) in Escherichia coli. The C fragment

CC comprises the C-terminal portion of the type A neurotoxin heavy

CC chain. It is pref. produced as a fusion protein with a

CC polyhistidine affinity tag or maltose binding protein, and is

CC used to produce avian antitoxin against C. botulinum type A

CC or as an immunogen in vaccine compsns. (see also AAR95009).

XX

SO Sequence 438 AA;

Query Match 100.0%; Score 727; DB 17; Length 438;
 Best Local Similarity 100.0%; Pred. No. 6.8e-77;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNS\$LRGT\$FTI\$K\$YAS\$G\$K\$D\$N\$V\$R\$N\$D\$V\$Y\$T\$N\$V\$V\$K\$N\$K\$E\$Y\$R\$L\$A\$T\$N\$A\$S\$Q\$A\$G\$E\$K\$I\$S\$A\$L 60

DB 292 LNS\$LRGT\$FTI\$K\$YAS\$G\$K\$D\$N\$V\$R\$N\$D\$V\$Y\$T\$N\$V\$V\$K\$N\$K\$E\$Y\$R\$L\$A\$T\$N\$A\$S\$Q\$A\$G\$E\$K\$I\$S\$A\$L 351

QY 61 EIPDVGNLSQVYVVKSKNDGINTKCKMNLQDNNNGNDIGFIGHQFNNTAKLVA\$MWYNR 120

DB 352 EIPDVGNLSQVYVVKSKNDGINTKCKMNLQDNNNGNDIGFIGHQFNNTAKLVA\$MWYNR 411

QY 121 QIERSSRTLGC\$SWEFI\$PVD 140

|||||

DB 412 QIERSSRTLGC\$SWEFI\$PVD 431

RESULT 10

AAW68389

ID AAW68389 standard; Protein; 438 AA.

XX AAW68389;

XX 07-DEC-1998 (first entry)

DE Clostridium botulinum toxin A fragment C.

XX Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

KW botulism.

OS Clostridium botulinum serotype A.

XX Key Location/Qualifiers

FT Misc-difference 1..2

FT note= "vector-derived amino acid residues"

PN WO9808540-A1.

PD 05-MAR-1998.

PF 28-AUG-1997; 97WO-US15394.

PR 28-AUG-1996; 96US-0704159.

PA (OPHI-) OPHIDIAN PHARM INC.;

PI Thalley BS, Williams JA;

DR WPI: 1998-230234/20.

DR N-PSDB; AAV30571.

XX Host cell containing recombinant expression vector encoding

PT Clostridium botulinum type B or E toxin - useful to treat humans

PT and other animals at risk of intoxication with clostridial toxin

XX Example 22; Page 262-263; 428pp; English.

XX This is the amino acid sequence of Clostridium botulinum serotype A

CC toxin C-fragment expressed by a DNA sequence (see AAV30571) in plasmid

CC PalterBot. Recombinant C-fragment proteins have been produced in

CC Escherichia coli as fusion proteins with either maltose binding

CC protein or Clostridium difficile type A toxin (see AAW68387). The

CC toxin relates to recombinant proteins derived from C. botulinum

CC toxins. Methods are provided which allow for the isolation of

CC soluble recombinant toxin proteins free of significant endotoxin

CC contamination. Preferred hosts for production of the recombinant

CC proteins are E. coli, insect cells and yeast cells. The recombinant

CC toxin proteins are used as immunogens for the production of vaccines

CC and antitoxins that are useful in the treatment of humans and

CC animals at risk of intoxication with clostridial toxin.

XX

Query Match 100.0%; Score 727; DB 19; Length 438;
 Best Local Similarity 100.0%; Pred. No. 6.8e-77;
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNS\$LRGT\$FTI\$K\$YAS\$G\$K\$D\$N\$V\$R\$N\$D\$V\$Y\$T\$N\$V\$V\$K\$N\$K\$E\$Y\$R\$L\$A\$T\$N\$A\$S\$Q\$A\$G\$E\$K\$I\$S\$A\$L 60

DB 292 LNS\$LRGT\$FTI\$K\$YAS\$G\$K\$D\$N\$V\$R\$N\$D\$V\$Y\$T\$N\$V\$V\$K\$N\$K\$E\$Y\$R\$L\$A\$T\$N\$A\$S\$Q\$A\$G\$E\$K\$I\$S\$A\$L 351

QY 61 EIPDVGNLSQVYVVKSKNDGINTKCKMNLQDNNNGNDIGFIGHQFNNTAKLVA\$MWYNR 120

DB 352 EIPDVGNLSQVYVVKSKNDGINTKCKMNLQDNNNGNDIGFIGHQFNNTAKLVA\$MWYNR 411

QY 121 QIERSSRTLGC\$SWEFI\$PVD 140

|||||

DB 412 QIERSRRTLGCSWEFIIPVD 431

RESULT 11

ID AAY77134

DB AAY77134 standard; Protein; 438 AA.

XX AAY77134;

XX 08-MAY-2000 (first entry)

XX Synthetic botulinum neurotoxin serotype A (BoNTA) C-terminal fragment.

DE Botulinum neurotoxin; heavy chain; BoNT; serotype A;

KM C-terminal fragment; Venezuelan equine encephalitis virus replicon;

KW VEE; botulism; vaccine; diagnosis; drug screening.

XX Clostridium botulinum.

OS Synthetic.

OS WO200002524-A2.

PN 20-JAN-2000.

XX 09-JUL-1999; 99WO-US15570.

XX 10-JUL-1998; 98US-0092416.

PR 12-MAY-1999; 99US-0133870.

XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.

XX Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;

PI WPI: 2000-160827/14.

XX N-PSDB; AA87212.

DR Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum

PT toxin serotypes A-G, is used for inducing an immune response against

PR botulinum -

XX Claim 22; Page 54; 54pp; English.

PS The invention relates to novel vaccines that induce a protective immune

CC response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F

CC and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant

CC DNA construct comprising a vector, and at least one nucleic acid

CC fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT

CC serotypes A-G. In preferred embodiments of the invention, the vector is

CC a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of

CC this vector results in the production of large amounts of a protein

CC encoded by a sequence cloned into the replicon. The constructs are used

CC to produce vaccines against botulism. The proteins can also be used as

CC diagnostic tools for the diagnosis of botulism. The transformed host

CC cells can be used to analyse the effectiveness of drugs and agents which

CC inhibit toxin effects. The vaccine currently used against botulism is

CC dangerous and expensive to produce, and contains formalin, which is very

CC painful for the recipient. Also, the vaccine is incomplete, in that only

CC 5 of the 7 serotypes are represented in the formulation. The novel

CC vaccine overcomes these problems, as it is easily purified, and

CC available in large quantities. It is also expressed in the lymph nodes

CC for a better immune response. Sequences AAY77134-Y77139 represent

CC synthetic BoNT Hc fragments used in the present invention. The DNA

CC encoding these sequences had been optimised for codon usage for

CC expression in yeast.

XX Sequence 438 AA:

SO Query Match 100.0%; Score 727; DB 21; Length 438;

Best Local Similarity 100.0%; Pred. No. 6,8e-77;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNSSLYRGTKFLIKKRYASGNKDNIVNRNDVYINVVKKKEYRLATNASQAGVEKILSAL 60

|||||

DB 292 LNSSLYRGTKFLIKKRYASGNKDNIVNRNDVYINVVKKKEYRLATNASQAGVEKILSAL 351

OY 61 EIPDVGNTLSOVVYKSKNOGINTNCKMNLQDNNGNDIGFIPHONNTAKLVASWYNR 120

DB 352 EIPDVGNTLSOVVYKSKNOGINTNCKMNLQDNNGNDIGFIPHONNTAKLVASWYNR 411

OY 121 QIERSRRTLGCSWEFIIPVD 140

DB 412 QIERSRRTLGCSWEFIIPVD 431

RESULT 12

AAW68391

ID AAW68391 standard; Protein; 445 AA.

XX AAW68391;

XX 07-DEC-1998 (first entry)

XX Clostridium botulinum toxin A fragment C (His-tagged).

DE Clostridium botulinum toxin A fragment C (His-tagged).

KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

KW botulism.

XX Clostridium botulinum serotype A.

OS Synthetic.

OS key Location/Qualifiers

FT Peptide 1..7

FT /note= "N-terminal histidine tag"

PN WO9808540-A1.

XX 05-MAR-1998.

XX 28-AUG-1997; 97WO-US15394.

XX 28-AUG-1996; 96US-0704159.

PR (OPHI-) OPHIDIAN PHARM INC.

XX (OPHI-) OPHIDIAN PHARM INC.

PI Thalley BS, Williams JA;

DR WPI: 1998-230234/20.

DR N-PSDB; AAV30576.

XX Host cell containing recombinant expression vector encoding

PT Clostridium botulinum type B or E toxin - useful to treat humans

PR and other animals at risk of intoxication with clostridial toxin

XX Example 29; Page 279-281; 428pp; English.

PS This is the amino acid sequence of a histidine-tagged fragment C

CC polypeptide of Clostridium botulinum serotype A toxin encoded by a

CC DNA sequence (see AAV30576) in plasmid pHisBot(syn). This vector

CC was used to express native soluble C fragment in *Escherichia coli*

CC host cells, with the recombinant C fragment being purified on a

CC poly-histidine binding affinity resin. The invention relates

CC to recombinant proteins derived from C. botulinum toxins. Methods

CC are provided which allow the isolation of soluble recombinant

CC proteins that are free of significant endotoxin contamination.

CC Preferred hosts for production of recombinant proteins are *E. coli*,

CC insect cells and yeast cells. The recombinant toxins are used as

CC immunogens for the production of vaccines and antitoxins that are

CC useful in the treatment of humans and animals at risk of

CC intoxication with clostridial toxin.

XX Sequence 445 AA:

SO Query Match 100.0%; Score 727; DB 19; Length 445;

Best Local Similarity 100.0%; Pred. No. 7e-77;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNSSLYRGTFEIIKKYASGKNKDNIVRNNDRVYINVVYKNEYRLATNASQAGVEKILISAL 60
 DB 299 LNSSLYRGTFEIIKKYASGKNKDNIVRNNDRVYINVVYKNEYRLATNASQAGVEKILISAL 358
 QY 61 EIPDVGNLSQVYVWYKSKNDGKITNCKKMNLDNNNGNDIGFIFGHQFNINIAKLVASMNYNR 120
 DB 359 EIPDVGNLSQVYVWYKSKNDGKITNCKKMNLDNNNGNDIGFIFGHQFNINIAKLVASMNYNR 418
 QY 121 QIERSSRTLGCSEFEIPVD 140
 DB 419 QIERSSRTLGCSEFEIPVD 438

RESULT 13

AAR95009 standard; Protein; 462 AA.

AC AAR95009;

DT 07-JUL-1996 (first entry)

DE Type A neurotoxin C fragment-polyhistidine tag fusion phisBot.

KW Toxin; neurotoxin; fusion protein; antitoxin; vaccine; immunogen;

KW Clostridium botulinum; polyhistidine; vector; PETHisa; phisBot.

OS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21 /label= Polyhistidine_tag

FT Protein 22..462 /label= C-fragment

PN W09612802-A1.

PD 02-MAY-1996.

PF 23-OCT-1995; 95WO-US13737.

PR 07-JUN-1995; 95US-0480604.

PR 24-OCT-1994; 94US-0329154.

PR 16-MAR-1995; 95US-0405496.

PR 14-APR-1995; 95US-0422711.

PA (OPHI-) OPHIDIAN PHARM INC.

PI Firca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;

PI Williams JA;

DR WPI; 1996-230603/23.

DR N-PADB; AAT929246.

PT Fusion proteins comprising non-toxin protein and part of toxin

PT useful to form anti-toxins against Clostridium botulinum type A, and

PT C. difficile type toxins, and to treat C. difficile intoxication,

PT partic. diarrhoea

PS Claim 7; Page 340-342; 434pp; English.

XX phisBot fusion protein (AAR95009), the product of a nucleotide

CC sequence (AAT929246) in vector PETHisa, comprises a polyhistidine

CC affinity tag and fragment C (see also AAR95008) of the Clostridium

CC botulinum type A neurotoxin. The phisBot protein was expressed

CC in Escherichia coli as a soluble protein and was purified by

CC metal chelate affinity chromatography to obtain a product free

CC of endotoxin contamination that may be useful as an immunogen

CC in vaccine compsns.

XX Sequence 462 AA;

QY Query Match 100.0%; Score 727; DB 17; Length 462;

Best Local Similarity 100.0%; Pred. No. 7.4e-77;

Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LNSSLYRGTFEIIKKYASGKNKDNIVRNNDRVYINVVYKNEYRLATNASQAGVEKILISAL 60
 DB 316 LNSSLYRGTFEIIKKYASGKNKDNIVRNNDRVYINVVYKNEYRLATNASQAGVEKILISAL 375
 QY 61 EIPDVGNLSQVYVWYKSKNDGKITNCKKMNLDNNNGNDIGFIFGHQFNINIAKLVASMNYNR 120
 DB 376 EIPDVGNLSQVYVWYKSKNDGKITNCKKMNLDNNNGNDIGFIFGHQFNINIAKLVASMNYNR 435
 QY 121 QIERSSRTLGCSEFEIPVD 140
 DB 436 QIERSSRTLGCSEFEIPVD 455

RESULT 14

AAW68390 standard; Protein; 462 AA.

AC AAW68390;

DT 07-DEC-1998 (first entry)

DE Clostridium botulinum toxin A fragment C (His-tagged).

KW Antitoxin; vaccine; neurotoxin; toxin A; intoxication; immunogen;

KW botulism.

OS Clostridium botulinum serotype A.

OS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21 /note= "N-terminal histidine tag"

PN W09808540-A1.

PD 05-MAR-1998.

PF 28-AUG-1997; 97WO-US15394.

PR 28-AUG-1996; 96US-0704159.

PA (OPHI-) OPHIDIAN PHARM INC.

PI Thalley BS, Williams JA;

PI WPI; 1998-230234/20.

DR N-PADB; AAV30572 and AAV30575.

PT Host cell containing recombinant expression vector encoding

PT Clostridium botulinum type B or E toxin - useful to treat humans

PT and other animals at risk of intoxication with clostridial toxin

PS Example 24; Page 265-267; 428pp; English.

XX This is the amino acid sequence of a histidine-tagged fragment C

CC polypeptide of Clostridium botulinum serotype A toxin encoded by a

CC DNA sequence (see AAV30572) in plasmid phisBot, and by a DNA sequence

CC (see AAV30575) in phisBotA. These vectors were used to express

CC native (i.e. non-fusion) soluble C fragment in Escherichia coli

CC host cells, with the recombinant C fragment being purified on a

CC poly-histidine binding affinity resin. The invention relates

CC to recombinant proteins derived from C. botulinum toxins. Methods

CC are provided which allow the isolation of soluble recombinant

CC proteins that are free of significant endotoxin contamination.

CC Preferred hosts for production of recombinant proteins are E. coli,

CC insect cells and yeast cells. The recombinant toxins are used as

CC immunogens for the production of vaccines and antitoxins that are

CC useful in the treatment of humans and animals at risk of

CC intoxication with clostridial toxin.

XX Sequence 462 AA;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:36:37 ; Search time 6.4553 Seconds
(without alignments)
899.518 Million cell updates/sec

Title: US-09-917-791-22

Perfect score: 727

Sequence: 1 LNSSLYRGTKEFIKKYASGN.....OIERSSRTLCGCSWEFIPYDD 140

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	727	100.0	1295	1 BXA1_CLOBO	P10845 clostridium
2	671	92.3	1295	1 BXA2_CLOBO	Q45894 clostridium
3	258.5	35.6	1274	1 BXF_CLOBO	P30996 clostridium
4	199	27.4	1250	1 BXE_CLOBO	Q00496 clostridium
5	190	26.1	1250	1 BXR_CLOBO	P30995 clostridium
6	165.5	22.8	1290	1 BXB_CLOBO	P10844 clostridium
7	163	22.4	1314	1 TERX_CLOBO	P04958 clostridium
8	158.5	21.8	1296	1 BXC_CLOBO	Q60393 clostridium
9	100	13.8	1290	1 BXC1_CLOBO	P18640 clostridium
10	88.5	12.2	264	1 MURI_FUSNN	O87666 fusobacteri
11	79.5	10.9	1786	1 YCPL_ARATH	P56785 arabidopsis
12	79	10.9	765	1 Y008_HUMAN	Q15398 homo sapien
13	75	10.3	720	1 WGC_ECOLI	P76387 escherichia
14	74	10.2	1276	1 BXD_CLOBO	P19321 clostridium
15	73.5	10.1	682	1 NISP_LACIA	Q07596 lactococcus
16	73	10.0	556	1 MUTL_SYNY3	P37349 synechocyst
17	73	10.0	573	1 TLPC_BACSU	P33209 bacillus su
18	73	10.0	720	1 WGC_ECO57	O84719 escherichia
19	72.5	10.0	875	1 ZIPL_YEAST	P31111 saccharomyc
20	72	9.9	1177	1 Y307_MYCEB	P47549 mycoplasma
21	71.5	9.8	399	1 L0LQ_BUCAI	P57382 buchnera ap
22	71.5	9.8	721	1 PRTP_HSVJ1	P53385 human herpe
23	71.5	9.8	735	1 CIGB_DICDI	O94481 dictyosteli
24	71.5	9.8	1076	1 RPOB_ASTIO	P27059 astasia ion
25	71.5	9.8	1744	1 TANA_XENIA	Q01550 xenopus lae
26	71.5	9.8	2339	1 RPCI_PLAFA	P27306 buchnera ap
27	71	9.8	266	1 TERM_BPPI2	P03681 bacterioph
28	71	9.8	353	1 RPOA_MYCSP	P38018 mycoplasma
29	71	9.8	445	1 IT2A_MOUSE	P15091 mus musculu
30	71	9.8	445	1 IT2B_MOUSE	Q97002 mus musculu
31	71	9.8	542	1 V155_FOWPV	Q915a7 fowlpox vir
32	71	9.8	3079	1 IRA2_YEAST	P19158 saccharomyc
33	71	9.8			

ALIGNMENTS

RESULT 1	ID	Sequence	Standard	PRT: 1295 AA.
AC	1 BXA1_CLOBO	P10845; P18639; P01561;		
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)			
DE	chain; Botulinum neurotoxin A (BOTOX) [Contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].			
GN	BOTA OR BNA OR ATX.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.			
OX	NCBI_TaxID=1491;			
RN	[1]			
RD	SEQUENCE FROM N.A.			
RE	STRAIN=NCCT 2916;			
RE	MEDLINE=90235864; PubMed=2185020;			
RA	Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.-J.,			
RA	Shone C.C., Atkinson T., Mellings J., Milton N.P.;			
RT	"The complete amino acid sequence of the Clostridium botulinum type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene."			
RT	Eur. J. Biochem. 189:73-81(1990).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=62A.			
RC	MEDLINE=97016817; PubMed=8863443;			
RA	East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;			
RT	"Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component."			
RT	Int. J. Syst. Bacteriol. 46:1105-1112(1996).			
RL	[4]			
RP	SEQUENCE OF 1-34 FROM N.A.			
RC	STRAIN=Hall;			
RC	MEDLINE=89350959; PubMed=2669749;			
RA	Betley M.J., Somers E., Dasgupta B.R.;			
RT	"Characterization of botulinum type A neurotoxin gene: delineation of the N-terminal encoding region."			
RL	Biochem. Biophys. Res. Commun. 162:1388-1395(1989).			
RN	[5]			
RP	SEQUENCE OF 1-18 FROM N.A.			
RC	STRAIN=Type A NIH;			
RC	MEDLINE=96096783; PubMed=8521962;			
RA	Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;			

004632 saccharomyc
092c28 rickettsia
P32571 saccharomyc
084877 chlamydia t
P38798 saccharomyc
P32051 escherichia
099p19 ureaplasma
085061 buchnera ap
P28289 homo sapien
P29767 clostridium
P34036 dictyosteli
O32481 legionella

RT "molecular characterization of two forms of nontoxic-nonhemagglutinating components of Clostridium botulinum type A progenitor toxins.";
RL FEBS Lett. 376:41-44(1995).
RN [6]
RP SEQUENCE OF 1-16.
RX MEDLINE=44178501; PubMed=6370252;
RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R. ;
RT "Partial amino acid sequence of the heavy and light chains of
RL botulinum neurotoxin type A.";
RN Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RA Dasgupta B.R., Foley J., Niece R. ;
RT "Partial sequence of the light chain of botulinum neurotoxin type A."
RL Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE=91120847; PubMed=21262206;
RA Dasgupta B.R., Dekleva M.L. ;
RT "Botulinum neurotoxin type A: sequence of amino acids at the
RL N-terminus and around the nicking site.";
RN Biochimie 72:661-664(1990).
RN [9]
RP SEQUENCE OF 448-464 AND 872-895.
RX MEDLINE=89024662; PubMed=31178218;
RA Sartymoorthy V., Dasgupta B.R., Foley J., Niece R.L. ;
RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
RL halves and their partial sequences.";
RN Arch. Biochem. Biophys. 266:142-151(1988).
RN [10]
RP SEQUENCE OF 448-482.
RX MEDLINE=85285016; PubMed=3896784;
RA Stone C.C., Hambleton P., Melling J. ;
RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
RL and purification of two tryptic fragments. Proteolytic action near
RT the COOH-terminus of the heavy subunit destroys toxin-binding
RT activity.";
RL Eur. J. Biochem. 151:75-82(1985).
RN [11]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94063091; PubMed=8243676;
RA Scienafio G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA Benfenati F., Wilson M.C., Montecucco C. ;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RL COOH-terminal peptide bonds.";
RN FEBS Lett. 335:99-103(1993).
RN [12]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=8294407;
RA Bluz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA Janin R., Nleman H. ;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
RN [13]
RP MUTAGENESIS OF GLU-261, PHE-265 AND TYR-365.
RX MEDLINE=21356941; PubMed=111700044;
RA Rigoni M., Cacchin P., Johnson E.A., Montecucco C., Rossetto O. ;
RT "Site-directed mutagenesis identifies active-site residues of the
RL light chain of botulinum neurotoxin type a.";
RN Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE=98455071; PubMed=9783750;
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C. ;
RT "Crystal structure of botulinum neurotoxin type A and implications
RL for toxicity.";
RN Nat. Struct. Biol. 5:898-902(1998).
RN [15]
RP FUNCTION: Inhibits acetylcholine release. The botulinum toxin
RL binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC

CC	-1	CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No detected action on small molecule substrates.
CC	-1	SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H).
CC	-1	SUBCELLULAR LOCATION: Secreted.
CC	-1	PHARMACEUTICAL: Available under the name BOTOX (Allergan) for the treatment of strabismus and blepharospasm associated with dystonia and cervical dystonia. Also used for the treatment of hemifacial spasm and a number of other neurological disorders characterized by abnormal muscle contraction.
CC	-1	MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC	-1	SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC	-1	DATABASE: NAME=BOTOX product information Web site; WWW="http://www.botox.com/index.jsp#productinfo".
CC	-1	DATABASE: NAME=Protein Spotlight; NOTE=issue 19 of February 2002;
CC		WWW="http://www.expasy.org/spottlight/articles/sp1t019.html".
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL:	X52066; CAA36289.1; -
DR	EMBL:	M30196; AAA23262.1; -
DR	EMBL:	X92973; CAA63551.1; -
DR	EMBL:	D67030; BAA1051.1; -
DR	EMBL:	M27892; AAA23269.1; -
DR	PIR:	A35294; BTCLAB.
DR	PIR:	S09492; S09492.
DR	PDB:	3BTA; O1-OCT-99.
DR	MEROPS:	M27_002; -
DR	InterPro:	IPR000395; Bontoxilysin.
DR	InterPro:	IPR000130; Zn_MTPeptide.
DR	Pfam:	PF01742; Peptidase_M27; 1.
DR	PRINTS:	PR00760; BONTOXILYSIN.
DR	Prodrom:	PD001963; Bontoxilysin; 1.
DR	PROSITE:	PS00142; ZINC_PROTEASE; 1.
KW	Neurotoxin;	Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW	Pharmaceutical;	3D-structure.
KW	INT_MET	0
FT	CHAIN	1 447
FT	CHAIN	448 1295
FT	METAL	222 222
FT	ACT_SITE	223 223
FT	METAL	226 226
FT	METAL	261 261
FT	DISULFID	429 453
FT	DISULFID	1234 1279
FT	TRANSMEM	626 646
FT	TRANSMEM	655 675
FT	VARIANT	26 26
FT	VARIANT	261 261
FT	MUTAGEN	265 265
FT	MUTAGEN	365 365
FT	CONFLICT	1 1
FT	CONFLICT	479 479
FT	CONFLICT	875 875
FT	CONFLICT	891 891
SO	SEQUENCE	1295 AA; 149322 MW; 856342f754862579 CRC64;

Query Match 100.0%; Score 727; DB 1; Length 1295;

Best Local Similarity 100.0%, Pred. No. 2.8e-59;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LNSSLYRGKFKFIKKYASGNKDNIVRNDRYINVVYKKEVRLATNNSOAGVEKILSL 60
DB 1149 LNSSLYRGKFKFIKKYASGNKDNIVRNDRYINVVYKKEVRLATNNSOAGVEKILSL 1208

OY 61 EIPDVGNLSQVYVYVMSKNDGITNCKMNLQDNGNDIGFIFGHFPNNIAKLVAASWNYR 120
DB 1209 EIPDVGNLSQVYVYVMSKNDGITNCKMNLQDNGNDIGFIFGHFPNNIAKLVAASWNYR 1268

OY 121 QIERSRRTGCSWEFIPVD 140
DB 1269 QIERSRRTGCSWEFIPVD 1288

RESULT 2
BXA2_CLOBO STANDARD; PRT: 1295 AA.
AC Q45894; P77780;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (BC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RA Williams A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins.";
RT other clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN-Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the RT gene encoding the nontoxic nonhemagglutinin component.";
RT Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin binds with high affinity to peripheral neuronal presynaptic membrane, is then internalized by receptor-mediated endocytosis. The C-terminus of the heavy chain (H) is responsible for the adherence of the toxin to the cell surface while the N-terminus mediates transport of the light chain from the endocytic vesicle to the cytosol. After translocation, the light chain (L) hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking neurotransmitter release. Inhibition of acetylcholine release results in flaccid paralysis, with frequent heart or respiratory failure (by similarity).
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H) (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----
CC DR EMBL: X73423; CAA51824.1; -;
DB EMBL: X87974; CAA61234.1; -;
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -;
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 447
FT CHAIN 448 1295
FT METAL 222 222
FT ACT_SITE 223 223
FT METAL 226 226
FT DISULFID 429 453
FT DISULFID 1234 1279
FT TRANSMEM 626 646
FT TRANSMEM 655 675
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;
Query Match 92.3%; Score 671; DB 1; Length 1295;
Best Local Similarity 90.7%; Pred. No. 4.1e-54;
Matches 127; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

OY 1 LNSSLYRGKFKFIKKYASGNKDNIVRNDRYINVVYKKEVRLATNNSOAGVEKILSL 60
DB 1149 LNSTLEGKFKFIKKYASGNEDNIVRNDRYINVVYKKEVRLATNNSOAGVEKILSL 1208

OY 61 EIPDVGNLSQVYVYVMSKNDGITNCKMNLQDNGNDIGFIFGHFPNNIAKLVAASWNYR 120
DB 1209 EIPDVGNLSQVYVYVMSKNDGITNCKMNLQDNGNDIGFIFGHFPNNIAKLVAASWNYR 1268

OY 121 QIERSRRTGCSWEFIPVD 140
DB 1269 QVGRASRTGCSWEFIPVD 1288

RESULT 3
BXF_CLOBO STANDARD; PRT: 1274 AA.
AC P30996;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (BC 3.4.24.69) (BONT/F)
DE (Bontoxilysin F).
GN BOTF.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson F.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.";
RT FEWS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN-Hobbs FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;

RX MEDLINE=94124495; PubMed=8294407;
 RA Jahn T., Blaszi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.,
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.",
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-1-ILE-
 CC 181 BOND IN SNAP-25.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62089; CAA43999.1; -;
 DR EMBL: X62583; CAA44558.1; -;
 DR PIR: A60027; A60027.
 DR PIR: B35294; B35294.
 DR PIR: JH0257; JH0257.
 DR PIR: S08575; S08575.
 DR PIR: S18111; S18111.
 DR PIR: S21178; S21178.
 DR HSSP: F10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mpeptidase.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 212 212 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 176 176 C -> G (IN REF. 2).
 FT CONFLICT 197 197 R -> S (IN REF. 2 AND 3).
 FT CONFLICT 339 339 R -> A (IN REF. 2).
 FT CONFLICT 772 772 I -> L (IN REF. 2).
 FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
 FT CONFLICT 966 966 R -> A (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 FT SEQUENCE 1250 AA; 143712 MW; D9FE2DDDA041EB4 CRC64;
 SQ
 Query Match 27.4%; Score 199; DB 1; Length 1250;
 Best Local Similarity 37.4%; Pred. No. 1.3e-10;
 Matches 52; Conservative 19; Mismatches 52; Indels 16; Gaps 6;
 Oy 1 LNSSLYRGTKFKIKKY-ASGNKDNIVRNNDRIYIVVY-KNEYVLAINASAGYEKILS 58
 Db 1116 LANRLYSIKYKIQVNNSSINDNLVRNDQVYINFAVSKTHLPLVADTAITNKEK--- 1172

Oy 59 ALEIPDVGN-LSQVYVYMKSKNDQGITNCKKNIADPNNGNDIGFGEHQFNNTAKLVASNW 117
 Db 1173 TIKISSSGKRFQVYVYNN-----VGNCTMAFKNKNNNGNIIIGF-----KADIVYASTW 1222
 Oy 118 VNRQIERSRTTLCGSWEFT 136
 Db 1223 YTHNRDHTNSNGCWFNFT 1241
 RESULT 5
 BSE.CLOBU
 ID BSE.CLOBU STANDARD; PRT; 1250 AA.
 AC P30995;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (BC 3.4.24.69) (BONT/E)
 DE (Bontoxilysin E).
 OS Clostridium/butyricum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 CC Clostridium.
 OX NCBI_TaxID=1492;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 43181, and ATCC 43755;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RL [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC STRAIN=BL6340;
 RX MEDLINE=91237316; PubMed=2033376;
 RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
 RA Yokosawa N., Yashiki T., Oguma K.;
 RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
 RT type E toxin gene from Clostridium butyricum strain BL6340";
 RL J. Gen. Microbiol. 137:519-525(1991).
 RL [3]
 RP SEQUENCE OF 1-48.
 RC STRAIN=5262;
 RA Gimenez J., Foley J., Dasgupta B.R.;
 RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
 RT partial sequence and comparison";
 RL FASEB J. 2:A1750-A1750(1988).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

DR EMBL: X62088; CAA43998.1; -
DR EMBL: X53180; CAA37321.1; -
DR PIR: JH0256; JH0256.
DR PIR: S16145; S16145.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT MET 0 0
FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 229 229 K -> M (IN REF. 2).
SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C2312857 CRC64;

Query Match 26.18; Score 190; DB 1; Length 1250;
Best Local Similarity 37.7%; Pred. No. 8.9e-10;
Matches 52; Conservative 19; Mismatches 53; Indels 14; Gaps 6;

QY 1 LNSLRGRKFLIKKY-ASGNKDNIVRNDRYINVVYKKEKRLATNMQAGVEKILSA 59
DB 1116 LARLKSIGVIGVYKRNSTNDLVKKNQVYINFA-SKTHLLPLVATATTKN-EKT 1173
QY 60 LEIPDGN-LSQVYVYKSKNDGITKCKKMLDNNNGNDIGFQFNIAVLVSNMY 118
DB 1174 IKSSSGNRNQVYVNS-----VGNCTNFKNNNGNNGIGLGF-----KADIVVASTW 1223
QY 119 NRQIERSSRTLGCSWEPT 136
DB 1224 YTHMRDNTNSNGFFNFI 1241

RESULT 6
BXB_CLOBO STANDARD; PRT; 1290 AA.
ID BXB_CLOBO STANDARD; PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
DE (Bontoxilysin B).
GN BOTB.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence."
RL Appl. Environ. Microbiol. 58:2345-2354(1992)...
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NCOTC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NCOTC 7273; PubMed=8408542;
RX MEDLINE=9401372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
```

```

RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin.";
RL Biochimie 70:811-817(1998).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Satyamurthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E.";
RL Arch. Biochem. Biophys. 238:544-548(1985).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RT "Botulinum neurotoxins are zinc proteins.";
RL J. Biol. Chem. 267:23479-23483(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RL Nature 359:832-835(1992).
RN [8]
RP FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RP RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
RP AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
RP WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
RP INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
RP ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
RP SYNAPTOSOMAL-2.
RN [9]
RP CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
RP neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
RP detected action on small molecule substrates.
RN [10]
RP SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
RP HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
RP WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
RP FORMATION AND TOXIN BINDING, RESPECTIVELY.
RN [11]
RP SUBCELLULAR LOCATION: Secreted.
RN [12]
RP MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
RP BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
RN [13]
RP SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
RN [14]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
```

DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR ProDom: PD001963; Bontotoxylysin: 1.
DR PROSITE, PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
FT METAL 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 230 230 BY SIMILARITY.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 436 445 INTERCHAIN (PROBABLE).
FT CONFLICT 29 29 T -> M (IN REF. 4).
FT CONFLICT 217 217 A -> G (IN REF. 2).
FT CONFLICT 224 224 S -> S (IN REF. 2).
FT CONFLICT 463 463 S -> R (IN REF. 4).
SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
Query Match 22.8%; Score 165.5; DB 1; Length 1290;
Best Local Similarity 28.7%; Pred. No. 1.7e-07;
Matches 43; Conservative 31; Mismatches 61; Indels 15; Gaps 5;
QY 5 LYRGTFIKKYYASGN--KDNIVRNNDVYINVVKKRYRLATNASQAGVEKLISALEI 62
DB 1138 LYIGKFTIRKRSNSQSIINDIVRKEDYIYDFENLQEMRYVYKYFKKEEKLFLAPI 1197
QY 63 PDVGLSGVVVVKSKNDGQIRKCKM--NLQDNNGNDIGFGFHQFNMA-----KL 112
DB 1198 SDSDEFYVTIQKEYDEQP--TYSQQLFKKDESDDELGLIGHFYESGIVFEYKDYF 1256
QY 113 VASNNYNNRQIERS--SPTLGSWEPIPYVD 140
DB 1257 CISKWKYKVKRKYNNLKLGCNMQPIPKDE 1286
RESULT 7
TETX.CLOTE
ID TETX.CLOTE STANDARD; PRT; 1314 AA.
AC P04958;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tetanus toxin precursor (Ec 3.4.24.68) (Tentoxylysin).
OS Clostridium tetani.
OG Plasmid.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RX SEQUENCE FROM N.A.
RA Eisel U., Jarausch W., Goretzki K., Henschen A., Engels J.,
RA Weller U., Hudel W., Habermann E., Niemann H.;
RA "Tetanus toxin: primary structure, expression in E. coli, and
RA homology with botulinum toxins";
RL EMBO J. 5:2495-2502(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CN3911;
RX MEDLINE=87040747; PubMed=3774547;
RA Fairweather N.F., Lyness V.A.;
RA "The complete nucleotide sequence of tetanus toxin";
RL Nucleic Acids Res. 14:7809-7812(1986).
RN [3]
RP SEQUENCE OF 742-1314 FROM N.A.
RX MEDLINE=86085672; PubMed=3510187;
RA Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;
RA "Cloning, nucleotide sequencing, and expression of tetanus toxin
RA fragment C in Escherichia coli";
RL J. Bacteriol. 165:21-27(1986).
RN [4]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=90201034; PubMed=2108021;

RA Kriegstein K., Henschen A., Weller U., Habermann E.;
RT "Arrangement of disulfide bridges and positions of sulfhydryl groups
RT in tetanus toxin";
RL Eur. J. Biochem. 186:39-45(1990).
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE=92037649; PubMed=1935979;
RA Kriegstein K.G., Henschen A.H., Weller U., Habermann E.;
RT "Limited proteolysis of tetanus toxin. Relation to activity and
RT identification of cleavage sites";
RL Eur. J. Biochem. 202:41-51(1991).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93010948; PubMed=1396558;
RA Schiavo G., Poullain B., Rossetto O., Benfenati F., Tauc L.,
RA Montecucco C.;
RT "Tetanus toxin is a zinc protein and its inhibition of
RT neurotransmitter release and protease activity depend on zinc";
RL EMBO J. 11:3577-3583(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93063293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poullain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin";
RL Nature 359:832-835(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.
RX MEDLINE=97475217; PubMed=9334741;
RA Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,
RA Sax M.;
RT "Structure of the receptor binding fragment HC of tetanus
RT neurotoxin";
RL Nat. Struct. Biol. 4:788-792(1997).
CC -I- FUNCTION: TETANUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 76-GLN-1-PHE-77
CC BOND OF SYNAPTOSOMAL-2.
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 76-GLN-1-PHE-77 BOND IN
CC SYNAPTOSOMAL-2.
CC -I- SUBUNIT: THE PRECURSOR POLYPEPTIDE IS SUBSEQUENTLY CLEAVED TO
CC YIELD SUBCHAINS L AND H. THESE REMAIN LINKED BY A DISULFIDE BRIDGE
CC AND ARE NON-TOXIC AFTER SEPARATION.
CC -I- MISCELLANEOUS: THE C-TERMINAL OF THE HEAVY CHAIN BINDS TO
CC GANGLIOSIDE RECEPTORS.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X04436; CAA28033.1; -;
DR EMBL; M12739; AAA23282.1; -;
DR EMBL; X06214; CAA29564.1; -;
DR PIR; A25689; BTCLTN;
DR PDB; 1AF9; 29-APR-98.
DR PDB; 1ABD; 14-OCT-98.
DR MEROPS; M27.001; -;
DR InterPro; IPR000395; Bontotoxylysin.
DR InterPro; IPR000130; zn_mtpeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontotoxylysin: 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc; Plasmid;

```

KW 3D-structure. 0
FT INIT_MER 1 456
FT CHAIN 457 1314 TETANUS TOXIN LIGHT CHAIN.
FT CHAIN 232 232 TETANUS TOXIN HEAVY CHAIN.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 233 233 BY SIMILARITY.
FT METAL 236 236 ZINC (CATALYTIC) (BY SIMILARITY).
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 669 689 POTENTIAL.
FT DISULFID 438 466 INTERCHAIN.
FT DISULFID 1076 1092
SQ SEQUENCE 1314 AA; 150550 MW; 134C3657133EF81D CRC64;

Query Match 22.4%; Score 163; DB 1; Length 1314;
Best Local Similarity 26.9%; Pred. No. 2.9e-07;
Matches 43; Conservative 26; Mismatches 49; Indels 42; Gaps 6;

QY 5 LKRGTFITIKKASGK-DNIYVANDRYIYINVKKEY-----RLA 45
DB 1168 LKNGLFITIKRYTPNNEIDSFVKSDFIKLYVSYNNNEHVGYPKDGNMFNNIDRLRVLG 1227
QY 46 TNASQAGVEKILSALEIPVGNLSQYVVKSKNDGITNCKKNLQDNNNGNDIGFIPGH-104
DB 1228 YNAPGIPLTKKMAVRLRDLK-----TISVQLKLYDKNKSGLVGTTH 1271
QY 105 -QFNNTAK--LIVASNMVNRQIERSSRTLCGSMFIPVDD 140
DB 1272 GQIGNDPNRDILASNMVFNHLK--DKILGCDMVFYPTDE 1309

RESULT 8
BXCL_CLOBO ID BXCL_CLOBO STANDARD; PRT; 1296 AA.
AC 060393:
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum
RT (Clostridium argentineense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins.";
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X74162; CAA52275.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INIT_MER 0
FT CHAIN 1 441
FT CHAIN 442 1296 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 230 230 BY SIMILARITY.
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F65C31 CRC64;

Query Match 21.8%; Score 158.5; DB 1; Length 1296;
Best Local Similarity 29.9%; Pred. No. 7.4e-07;
Matches 46; Conservative 28; Mismatches 59; Indels 21; Gaps 7;

QY 4 SLKRGTFITIKKASG---NKNIVANDRYIYIN-VVANKKEYRLATNASQAGVEKILSA 59
DB 1143 NYLGLRFLITIKKASNSRINNDINREGYIYLNIDNISDESRYVLVNSKEIQTLEL 1202
QY 60 LEIPDVGNLSQYVVKSKNDGITNCKKNLQDNNNGNDIGFIPGHOF-----NN 108
DB 1203 APIINDPTFYDVQLK-KYREKTYNCQI-LCEKDKRTGELGICGFVADGYVMDTYDN 1260
QY 109 IAKTVASNMVNRQIER--SSRTLCGSMFIPVDD 140
DB 1261 Y-FCISQWYLRIRISENINKRLRGCMWGFIPVDE 1292

RESULT 9
BXCL_CLOBO ID BXCL_CLOBO STANDARD; PRT; 1290 AA.
AC P18640:
DT 01-NOV-1990 (rel. 16, Created)
DT 01-NOV-1990 (rel. 16, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BONT/C1)
DE (Bontoxilysin C1).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=90370487; PubMed=2204031;
RX Hauser D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gili D.M.,
RA Boquet P., Popoff M.R.;
RT "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
RL Nucleic Acids Res. 18:4924-4924(1990).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=Type C Stockholm / C-ST;
CC MEDLINE=91024998; PubMed=2222445;
RX Kimura K., Fujii N., Tsuzuki K., Murakami T., Indoh T.,
RA Yokosawa N., Takeshi K., Syuto B., Oguma K.;
RT "The complete nucleotide sequence of the gene coding for botulinum
RT type C1 toxin in the C-ST phage genome.";
RL Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
CC [3]
CC SEQUENCE OF 2-25.

```

RC STRAIN-Type C Stockholm / C-ST;
 RX MEDLINE=88153072; PubMed=2450068;
 RA Tsubaki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
 RA Oguma K.,
 RT "Establishment of a monoclonal antibody recognizing an antigenic site
 RT common to Clostridium botulinum type B, C1, D, and E toxins and
 RT tetanus toxin.";
 RL Infect. Immun. 56:898-902(1988).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94038966; PubMed=7901002;
 RA Blaszi J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.,
 RT "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
 RT cleaving HPC-1/syntaxin.";
 RL EMO J. 12:4821-4828(1993).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES. IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES SYNTAXIN
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X66433; CAA47060.1; -;
 DR EMBL: X72793; CAA51313.1; -;
 DR EMBL: X53751; CAA37780.1; -;
 DR EMBL: D90210; BAA14235.1; -;
 DR EMBL: X62389; CAA44263.1; -;
 DR PIR: S11291; S11291.
 DR PIR: A35396; A35396.
 DR PIR: A43503; A43503.
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MPeptidase.
 DR Pfam: PF01742; Peptidase_M27.1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin.1.
 DR PROSITE: PS00142; ZINC_PROTEASE.1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; zinc.
 FT INT MET 0
 FT CHAIN 1 448 BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
 FT CHAIN 449 1290 BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACCT SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 452 INTERCHAIN (PROBABLE).
 FT CONFLICT 84 84 P -> T (IN REF. 2).
 SO SEQUENCE 1290 AA; 148734 MW; 71BE379F91129E8 CRC64;

Query Match 13.8%; Score 100; DB 1; Length 1290;
 Best Local Similarity 23.6%; Pred. No. 0.18;

Matches 37; Conservative 21; Mismatches 77; Indels 22; Gaps 4;
 QY 2 NSSLYRGFTFLTKKASGKNIVRNDRVYINVVYKNEYRL-----ATNSAQGV 53
 DB 1138 NDFNEGYIIIRKIRGTNTDVRGSDILYDPTNNKAYLFPKNETMYDNHSTEDI 1197
 QY 54 EKLSALEIPDVGNSQVYVMSKNDQGITNC-KMNLDDNGNDIGFIFHQF----- 106
 DB 1198 YALGLEQKTDINDNINIIQIDPMNNTYYASQIFKSNFNGEINSGICIGYRFLRGDW 1257
 QY 107 ---NNIAKLVASNWNROIRSSRFLGCSWEPIVD 140
 DB 1258 YRHNLYLPTVKOGNYASLLESTST---THMGVPVSE 1290
 RESULT 10
 MURI_FUSNN
 ID MURI_FUSNN STANDARD; PRT; 264 AA.
 AC Q8RE6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Glutamate racemase (EC 5.1.1.3).
 GN MURI OR FN1161.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacterium.
 OX NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 25586;
 RX MEDLINE=21886394; PubMed=11889109;
 RA Kaprielian V., Anderson I., Ivanova N., Reznik G., Los T., Lykdis A.,
 RA Bhattacharyya A., Bartman A., Gardner W., Gueckin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Golsman T., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Forstner M., Kyriades N., Overbeek R.,
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium
 RT nucleatum strain ATCC 25586.";
 RL J. Bacteriol. 184:2005-2018(2002).
 CC -1- FUNCTION: Provides the (R)-glutamic acid required for cell wall
 CC biosynthesis (by similarity).
 CC -1- CATALYTIC ACTIVITY: L-glutamate -> D-glutamate.
 CC -1- PATHWAY: Peptidoglycan biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTATE/GLUTAMATE RACEMASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AE010622; AAL95357.1; -;
 DR PROSITE: PS00923; ASP_GLU_RACEMASE.1;
 DR PROSITE: PS00924; ASP_GLU_RACEMASE.2;
 KW Peptidoglycan synthetase; Cell wall; Isomerase; Complete proteome.
 SO SEQUENCE 264 AA; 29882 MW; F69A0C16BBA8AF9 CRC64;

Query Match 12.2%; Score 88.5; DB 1; Length 264;
 Best Local Similarity 26.7%; Pred. No. 0.36;
 Matches 39; Conservative 22; Mismatches 54; Indels 31; Gaps 6;
 QY 16 YASGKNQIVRNDRVYINVVYK-N-KERLATN-ASQAGVKILSALEIPDVGNSOVY 72
 DB 44 YSGKTKMELQKTERILDFEFAKNNCKLVYACNTASTAIDYLRKRSPLTGITEAGV 103
 QY 73 VKSKRN-----DQGITNCKNMLDNGNDIGFIFHQF-----FNN 108
 DB 104 KIASKNTKKNKNAVISTKRTAESHGKNAKMLDSELNKEIACIEFPAQMIETGMDTFDN 163
 QY 109 IAKLVASNWNROIRSSRFLGCSWEPIVD 132
 :|: | | :|: :|: | | |

DB 164 KREL--NKYLSEIPKNDTLVIGCT 187

RESULT 11

YCF1_ARATH STANDARD; PRT; 1786 AA.

AC P56785;

DT 30-MAY-2000 (Rel. 39, Created)

DR 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 213.7 kDa protein ycf1.

GN YCF1

OS Arabidopsis thaliana (Mouse-ear cress).

OC Chloroplast.

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=20039611; PubMed=10574454;

RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;

RT "Complete structure of the chloroplast genome of Arabidopsis thaliana."

RL DNA Res. 6:283-290(1999).

CC -1- FUNCTION: NOT YET KNOWN.

CC -1- SIMILARITY: BELONGS TO THE YCF1 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: AP000423; BAA84445.1; -

DR EMBL: AP000423; BAA84433.1; -

KW Chloroplast; Hypothetical protein.

SQ SEQUENCE 1786 AA; 213727 MW; CFFD2A4D776D7E5D CRC64;

Query Match 10.9%; Score 79.5; DB 1; Length 1786;

Best Local Similarity 25.5%; Pred. No. 20;

Matches 40; Conservative 26; Mismatches 58; Indels 33; Gaps 8;

DB 1103 FETSKTILDKYIKNEENGKKNTLYFISTIKNLISNKKKMSYDLC-SLSQAYFYFKLS 1161

QY 59 ALEIPDVGNLSQVY-----VMKSK-----NDQGI-----TNKCKNKLQDNGNDIG 99

DB 1162 QIKVSNFKKAKAVLEYNICITSEFVKNKIKVFQEHGIFHYELKNTFLNSEVNQMKM-W 1220

QY 100 FIGFHQFN-----NIKLVASNMWYNRQIESKSLGCS 132

DB 1221 LRSQYQYNLPQISWARYLVQNMKNK-INKDSVLNPS 1256

RESULT 12

Y008_HUMAN STANDARD; PRT; 765 AA.

AC Q15398;

DT 15-JUL-1998 (Rel. 36, Created)

DR 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein KIAA0008.

GN KIAA0008.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=96051387; PubMed=7584026;

RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawabayashi Y.,

RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;

RT "Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1."

RL DNA Res. 1:27-35(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Eye, and Lung;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: D13633; BAA02797.1; -

DR EMBL: BC010658; AAH10658.1; -

DR EMBL: BC016276; AAH16276.1; -

DR InterPro: IPR005026; GKAP.

DR pfam: PF03359; GKAP; 1.

KW Hypothetical protein

SQ SEQUENCE 765 AA; 85668 MW; 00AFF91A0387EAL CRC64;

Query Match 10.9%; Score 79; DB 1; Length 765;

Best Local Similarity 21.9%; Pred. No. 8;

Matches 32; Conservative 27; Mismatches 39; Indels 48; Gaps 8;

QY 20 NKNIVNRNDRYIN-----VKNKRYLA-----TASQAGVEKILS--- 58

DB 309 HEEHVINKKEATYKNGLPKEVPSLERNGRIAPHGVPYFRN1LQSEKELTSHCF 368

QY 59 -----ALEIPD-----VGNLSQVYVMSKNOGKITNC--KMNLDNNGNDI-GF 100

DB 369 EMDRKLELDIPDAKDLIRAVGQTRLMKERKQFEGLVDCFYKRGKIKETCTLDIGF 428

QY 101 ---IGF-----HGFNNIAKLVASNM 117

DB 429 WDMVSFQIEDVYHKFNKLKLEESGW 454

RESULT 13

WZC_ECOLI STANDARD; PRT; 720 AA.

AC P76387; P71236; O08003; O08004;

DT 15-JUN-2002 (Rel. 41, Created)

DR 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tyrosine-protein kinase wzc (EC 2.7.1.112).

GN WZC OR B2060.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=96326333; PubMed=8759852;

RA "Stevenson G., Andrianopoulos K., Hobbs M., Reeves P.R.;

RT "Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid.";

RL J. Bacteriol. 178:4885-4893(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RA Sunagawa H., Ohyama T., Watanabe T., Inoue K.;
 RT "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phage d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN-D-SA, and D-1873;
 RX MEDLINE-89339741; PubMed-2668193;
 RA Morishik K., Syuto B., Kubo S., Oguma K.;
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RT strains.";
 RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE-94230352; PubMed-8175689;
 RA Yamasaki S., Baumeister A., Binz T., Blas J., Link E., Cornille F.,
 RT Rogues B., Fyfe E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT Cleavage of members of the synaptobrevin/VAMP family by types D and
 RL F botulin neurotoxins and tetanus toxin.";
 J. Biol. Chem. 269:12764-12772(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-1-LEU-61 BOND OF
 CC SYNAPTOSOMALIN-1 AND -2.
 CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (1) AND A
 CC HEAVY CHAIN (4). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X54254; CAA38175.1; -
 DR EMBL; S49407; AAB24244.1; -
 DR PIR: S11455; S11455.
 DR HSSP: P10845; 3BTA.
 DR MEROPS; M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptidse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT CHAIN 1 442
 FT CHAIN 1276
 FT METAL 229
 FT ACT_SITE 230
 FT METAL 233
 FT METAL 233
 FT DISULFID 437
 FT VARIANT 15
 FT VARIANT 17
 FT VARIANT 452
 FT VARIANT 457
 R -> Q (IN STRAIN D-SA).
 K -> T (IN STRAIN D-SA).
 R -> T (IN STRAIN D-SA).

FT VARIANT 457 457 R -> F (IN STRAIN D-1873).
 FT VARIANT 462 462 A -> D (IN STRAIN D-1873).
 FT VARIANT 489 489 K -> N (IN STRAIN CB16).
 FT VARIANT 644 644 N -> K (IN STRAIN CB16).
 FT VARIANT 1122 1122 O -> R (IN STRAIN CB16).
 FT SEQUENCE 1276 AA; 146871 MW; C1EC50F6C8233E2 CRC64;
 Query Match 10.2%; Score 74; DB 1; Length 1276;
 Best local Similarity 21.9%; Pred. No. 45;
 Matches 30; Conservative 18; Mismatches 39; Indels 50; Gaps 5;
 QY 3 SS1RGTKFLIKKYGASGNDIVRNDRYINVVANKEKRLATN-----ASQAGYEKIL 57
 DB 1127 SKLYTGNPTIKKSVSDKNPYSRILNDNIIHLMLNSRKRYMILRDITVATGG----- 1181
 QY 58 SALEIPDVGNLSQVVMKSNDDGINTCKMN-----LDDNNGN-DIGIFGHQFNNTA 110
 DB 1182 -----ECSQNCVYALKLQSNLGNVGIGIF-----SIK 1208
 QY 111 KLVASWYNRQIERSSR 127
 DB 1209 NIVSKNKYCQIFSSFR 1225
 RESULT 15
 NISP_LACIA STANDARD; PRT; 682 AA.
 AC 007596;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nisin leader peptide processing serine protease nisp precursor
 DE (EC 3.4.21.-).
 GN NISP.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIZO R5;
 RX MEDLINE-93239683; PubMed-8478324;
 RA van der Meer J.R., Polman J., Beerthuyzen M.M., Smeets R.J.,
 RT Kuipers O.P., de Vos W.M.;
 RT "Characterization of the Lactococcus lactis nisin A operon genes
 RT nisp, encoding a subtilisin-like serine protease involved in
 RT precursor processing, and nlsr, encoding a regulatory protein
 RT involved in nisin biosynthesis.";
 RL J. Bacteriol. 175:2578-2586(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-6F3;
 RX MEDLINE-94213458; PubMed-8161176;
 RA Engelke G., Gutowski-Eckel Z., Klesau P., Siegers K.,
 RA Hammelmann M., Ertlan K.-D.;
 RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
 RT 6F3.";
 RL Appl. Environ. Microbiol. 60:814-825(1994).
 RN [3]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-95357326; PubMed-7630881;
 RA Smeets R.J., Kollmeier H.S., Kuipers O.P., de Vos W.M.;
 RT "Homology modelling of the Lactococcus lactis leader peptidase Nisp
 RT and its interaction with the precursor of the antibacterial nisin.";
 RL Protein Eng. 8:117-125(1995).
 CC -1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
 CC -1- PATHWAY: Nisin biosynthesis; last step.
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (Potential).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)

DR EMBL: L11061; AAA25200.1; -
DR EMBL: X76884; CAA54210.1; -
DR HSSP: P29600; 1GCI.
DR MEROPS: S08.059; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; FALSE_NEG.
DR PROSITE: PS50847; GRAM_POS_ANCHORING; FALSE_NEG.
KM Hydrolyase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 195 POTENTIAL.
FT CHAIN 196 655 NISIN LEADER PEPTIDE PROCESSING SERINE
PROTEASE NISP.
FT PROPEP 656 682 REMOVED BY SORTASE (POTENTIAL).
FT ACT_SITE 259 259 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 306 306 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 512 512 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 652 656 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 655 655 AMIDE-LINKED TO CELL WALL (POTENTIAL).
FT CONFLICT 500 500 A -> T (IN REF. 2).
SQ SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Query Match 10.1%; Score 73.5; DB 1; Length 682;
Best Local Similarity 25.2%; Pred. No. 25;

Matches 37; Conservative 22; Mismatches 51; Indels 37; Gaps 8;

OY 12 IIRKYASGNKDNIVR-----NNDRYINVV-----YKNKEYRLATNASOAGYEKILS 58
DB 525 VYDKYGIKPNQKRLKRLNNSPEVNGNRV-LNIYDLNCKNKAFLDTDKGD-----D 577
OY 59 ALEIDVGNLSOV-VYMKSKNDGITNCKCKMNLQDNGNDIGFIGHQFNNTAKLVASNM 117
DB 578 AINHKSMENLKESRDTMKQEDKEIORNTNNNFSIKN-----DFHNISKEYISVD 627
OY 118 Y--NROIERSRLGC---SWEFIPV 138
DB 628 YNINQMANNNRNSRGAVSYRQEIILPV 654

Search completed: March 13, 2003, 11:40:38
Job time : 8.45533 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:31:47 ; Search time 20.9798 seconds
(without alignments)
1374.968 Million cell updates/sec

Title: US-09-917-791-22

Perfect score: 727

Sequence: 1 LNSSLYRGTRFKTIKKYASGN.....QIERSRRLGCSWEFIPYDD 140

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	341	46.9	72	2	Q9R540
2	282.5	38.9	1268	2	Q4S851
3	277.5	38.2	1278	2	Q57236
4	260.5	35.8	1280	2	Q9ZAJ5
5	205	28.2	1251	2	Q9K395
6	197.5	27.2	1255	2	Q9FAR6
7	169.5	23.3	1291	2	Q08077
8	162	22.3	451	2	Q9LAI3
9	161.5	22.2	441	2	Q9X708
10	160.5	22.1	1291	2	Q93G71
11	160.5	22.1	1291	2	Q933K0
12	160.5	22.1	1291	2	Q9ZAJ8
13	160	22.0	1310	2	Q93N27
14	110	15.1	315	16	Q97F01
15	100	13.8	1291	2	Q93HT3
16	95	13.1	1285	2	Q9LBR1

17	95	13.1	1285	2	Q4S967	Q4S967 clostridium
18	88.5	12.2	264	16	Q8RE6	Q8RE6 fuscobacteri
19	88	12.1	408	15	Q11572	Q11572 human immun
20	87.5	12.0	204	15	Q9DV80	Q9DV80 human immun
21	86.5	11.9	401	15	Q11575	Q11575 human immun
22	85.5	11.8	467	16	Q97N26	Q97N26 clostridium
23	85	11.7	1593	5	Q8T177	Q8T177 dictyosteli
24	84.5	11.6	848	15	Q9WC60	Q9WC60 human immun
25	83	11.4	245	16	Q8RDM8	Q8RDM8 fuscobacteri
26	83	11.4	408	15	Q11570	Q11570 human immun
27	83	11.4	409	15	Q11571	Q11571 human immun
28	83	11.4	1483	3	Q96UL6	Q96UL6 diaporthe a
29	83	11.4	4550	5	Q77336	Q77336 plasmodium
30	82.5	11.3	479	10	Q9M664	Q9M664 nicotiana t
31	82	11.3	255	16	Q8XMA0	Q8XMA0 clostridium
32	82	11.3	408	15	Q11567	Q11567 human immun
33	81.5	11.2	235	2	Q93D91	Q93D91 streptococc
34	81.5	11.2	408	15	Q11574	Q11574 human immun
35	81.5	11.2	1077	1	Q97704	Q97704 methanosarc
36	81.5	11.2	1869	12	Q8V2H2	Q8V2H2 camelipox vi
37	81.5	11.2	1869	12	Q8QPZ7	Q8QPZ7 camelipox vi
38	81	11.1	336	12	Q91M09	Q91M09 lumby skin
39	81	11.1	1301	5	Q8WSK5	Q8WSK5 plasmodium
40	81	11.1	1484	3	Q8TGV4	Q8TGV4 diaporthe p
41	80	11.0	1122	3	Q8X1C2	Q8X1C2 photopsis s
42	79.5	10.9	214	15	Q74058	Q74058 human immun
43	79.5	10.9	233	16	Q92C15	Q92C15 listeria in
44	79.5	10.9	307	16	Q8XK10	Q8XK10 clostridium
45	79.5	10.9	1807	5	Q8WSL1	Q8WSL1 plasmodium

ALIGNMENTS

RESULT 1						
Q9R540		PRELIMINARY:	PRT:	72 AA.		
AC	Q9R540:					
DT	01-MAY-2000 (TREMBLrel. 13, Created)					
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)					
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)					
DE	Neurotoxin heavy chain 18 kDa fragment (Fragment).					
OS	Clostridium botulinum.					
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;					
CC	Clostridiales; Clostridiaceae; Clostridium.					
OX	NCBI_TaxID=1491;					
RN	[1]					
RP	SEQUENCE.					
RP	MEDLINE=94000342; PubMed=8397793;					
RA	Gimenez J.A., Dasgupta B.R.;					
RT	"Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72,					
RT	45, 42, and 18 kD fragments.";					
RT	J. Protein Chem. 12:351-363(1993).					
DR	HSP; P10845; 3BPA.					
SQ	SEQUENCE 72 AA; 8165 MW; B7A959576A615E18 CRC64;					
Query Match	46.9%;	Score 341;	DB 2;	Length 72;		
Best Local Similarity	98.6%;	Pred. No. 1.7e-24;				
Matches 69;	Conservative	0;	Mismatches 1;	Indels	0;	Gaps 0;
OY	1 LNSSLYRGTRFKTIKKYASGNKDINVRNDYINVVKKKEYRLATNASQAVEKLSAL 60					
DB	3 LNSSLYRGTRFKTIKKYASGNKDINVRNDYINVVKKKEYRLATNASQAVEKLSAL 62					
OY	61 EIPDVGNLSQ 70					
DB	63 EIPDVGNLSQ 72					
RESULT 2						
ID	Q4S851	PRELIMINARY:	PRT:	1268 AA.		
AC	Q4S851:					

DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE Neurotoxin type F.
GN BONT /F.
OS Clostridium baratti.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
ON NCBI_Taxid=1561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93252228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium baratti type F
neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEMS Microbiol. Lett. 108:175-182(1993).
DR EMBL: X68262; CAA48329.1; -.
DR HSSP: P10845; 3B7A.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 38.9%; Score 282.5; DB 2; Length 1268;
Best Local Similarity 45.3%; Pred. No. 1.5e-17;
Matches 63; Conservative 19; Mismatches 44; Indels 13; Gaps 5;

QY 2 NSSLYGTFFIIRKVAS---GNKDNIVRNDRVYINVVVKNKRYLATNASQAVEKIIS 58
DB 1131 NARLYGVEVILIRKNGSTDSINTDNFVRKNDYIINVVQNSRYQLADVSTSAVEKTIK 1190
QY 59 ALEIFPVG-LSGVVYVVKSKNDGKITNCKMNIQDNNNGNDIGTIGHOENNIKLVASNW 117
DB 1191 LRRISNYSNNYSNMQIIMDS----IGDNCTMNFKNYNGNDIGLGFH-LNN--LVASSW 1241
QY 118 YNRQIERSRFTLCGSWEFI 136
DB 1242 YKNNIRNNTNNGCFSFI 1260

RESULT 3

ID 057236 PRELIMINARY; PRT: 1278 AA.
AC 057236; Q45863;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, last annotation update)
DE Botulinum neurotoxin type F (BONT/F protein).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
ON NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 10281.
RA Hutson R.A., Collins M.D.;
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 635-1000 FROM N.A.
RC STRAIN=NCTC 1028;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and

RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]

RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=LANCELAND;
RX MEDLINE=98404102; PubMed=9732534;
RA East A.K., Bhandari M., Hieim S., Collins M.D.;
RT "Analysis of the botulinum neurotoxin type F gene clusters in
RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT baratti";
RL Curr. Microbiol. 37:262-268(1998).
DR EMBL: X61714; CAA57358.1; -.
DR EMBL: L35496; AAA23210.1; -.
DR EMBL: X70821; CAA50152.1; -.
DR EMBL: X99064; CAA67512.1; -.
DR HSSP: P10845; 3B7A.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 38.2%; Score 277.5; DB 2; Length 1278;
Best Local Similarity 44.6%; Pred. No. 4.5e-17;
Matches 62; Conservative 19; Mismatches 45; Indels 13; Gaps 4;

QY 2 NSSLYGTFFIIRKVAS---GNKDNIVRNDRVYINVVVKNKRYLATNASQAVEKIIS 58
DB 1140 NTRLYGVEVILIRKNGSTDSINTDNFVRKNDLAYIINVVQNSRYQLADVSTSAVEKTIK 1199
QY 59 ALEIFPVG-LSGVVYVVKSKNDGKITNCKMNIQDNNNGNDIGTIGHOENNIKLVASNW 117
DB 1200 LIFTSNNSNLGQIIVWDS----IGNCTMNFQNNNGNDIGLGFHSNN--LVASSW 1250
QY 118 YNRQIERSRFTLCGSWEFI 136
DB 1251 YKNNIRKNTSSNCGFWSFI 1269

RESULT 4

ID 092A75 PRELIMINARY; PRT: 1280 AA.
AC 092A75;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE Bont protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
ON NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 3281;
RX MEDLINE=98440323; PubMed=9767710;
RA Santos-Bueiga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins";
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13631; CAA73972.1; -.
DR HSSP: P10845; 3B7A.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1280 AA; 147487 MW; D0F748976EBC222C CRC64;
Query Match 35.8%; Score 260.5; DB 2; Length 1280;
Best Local Similarity 42.9%; Pred. No. 1.7e-15;
Matches 60; Conservative 19; Mismatches 48; Indels 13; Gaps 4;
QY 1 LNSSLYRGTFILIKKYAS--GNKDNIVRNNDRIYINVVYKNEKRYLATNASQAGVEKIL 57
DB 1142 LNKYLEGEVILIRKNAPIDISWDFVRKNDLAYINVDHGYELLYADISTTKEKII 1201
QY 58 SALEIPDVG-NLSQVYVYKSKNDGITTNCCKNMLQDNNGNDIGFIFGHQFNNTAKLVASN 116
DB 1202 KLRTSNPNDSLGQIIVMDS-----IGNCTMNFQNNDSNIGLIGFHSDD---LVASS 1252
QY 117 WYNROIESSRTLCGSWEFI 136
DB 1253 WYNNHRRNTSSNGCFWFSFI 1272
RESULT 5
09K395
ID 09K395 PRELIMINARY; PRT; 1251 AA.
AC 09K395;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Type E botulinum toxin.
GN BONT/E.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 095;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 095) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 155;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
RA Kato H., Nakamura S., Karasawa T.;
RT "C. butyricum (LCL 155) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1899;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1897;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1898;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1886;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;

RT "C. butyricum (KZ 1886) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1887;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1889;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1889) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1890;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1890) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1891;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1891) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 063;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 063) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 095;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 095) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL 155;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (LCL 155) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [14]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1899;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1899) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [15]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1897;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1898;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;
RT "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
RN [17]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1886;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
RA Karasawa T.;

```

RESULT 6
O9FAR6 PRELIMINARY; PRT; 1255 AA.
AC O9FAR6;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Type E botulinum toxin.
GN BONT/E.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BL 6340/ATCC 43755/BL 5520/K2 147;
RX MEDLINE=20509829; PubMed=11055954;
RA Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y.,
RA Yamakawa K., Oguma K., Sakaguchi Y., Nakamura S.;
RT "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium
RT butyricum Strains.",
RL Appl. Environ. Microbiol. 66:4992-4997(2000).
DR HSSP; P10845; 3BTA.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1255 AA; 143918 MW; 1B557B9D85C8EAD CRC64;

Query Match 27.2%; Score 197.5; DB 2; Length 1255;
Best Local Similarity 38.4%; Pred. No. 1.2e-09;
Matches 53; Conservative 20; Mismatches 52; Indels 13; Gaps 6;

OY 1 LNSLYRGTEFIIRKY-ASGNKDNIVANNDRVYINVVYKKEKRLATNASQAGEKILSA 59
DB 1120 LANRLVSGIKVKTQVRVNNSTNDIVAKNDQYVINFPA-SKTHLLPLPYATFTATTNK-EKT 1177
OY 60 LEIPDVGN-LSQVYVMSKNDGKITNCKMNLDDNNGNDIGTGFHGFNNIAKLVASNMV 118
DB 1178 IKISSGGRNENQVYVYVMS-----VGNCTMFRNNGNINIGLGF---KADTVVASTWY 1228
OY 119 NRQIERSRRLGCSWERT 136
DB 1229 YTHMDNTNSGFFWMT 1246

RESULT 7
O08077 PRELIMINARY; PRT; 1291 AA.
AC O08077;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DE 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE Botulinum neurotoxin type B (EC 3.4.24.-) (BONT/B).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EKLDN 17B ATCC25765;
RX MEDLINE=94122659; PubMed=7764370;
RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic
RT Clostridium botulinum type B neurotoxin: comparison with other
RT Clostridia neurotoxins.",
RL Curr. Microbiol. 28:101-110(1994).

```

```

CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOBREVIN-2.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC NEUROTOXIN.
CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC REGION.
DR EMBL; X71343; CAA50482.1; -
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
SQ SEQUENCE 1291 AA; 150513 MW; 71BCAFEE23D65FAAA CRC64;

```

```

Query Match 23.3%; Score 169.5; DB 2; Length 1291;
Best Local Similarity 28.3%; Pred. No. 4.8e-07;
Matches 43; Conservative 39; Mismatches 53; Indels 17; Gaps 7;

```

```

OY 4 SLRYGRTEFIIRKYSAGN-KDNIVANNDRVYINVVYKKEKRL-ATNASQAGEKILSLA 60
DB 1138 NLTYGEKFIIRRESNOSINDIVRKEDYTHDLVHHEMRVAYAKYRKDEKELFLSI 1197
OY 61 EIPDVGNLSQVYVMSKNDGKITNCKM-NLDDNNGNDIGTGFHGFNNIAKL----- 112
DB 1198 -IDSNSMEFYETIEKEVDEBP-SYSCQLLEKKEDESPDDIGLGHFRFYSGLVLRKKYKD 1255
OY 113 --VASMNYNRQIERS--SRILGCSWERTIPVD 140
DB 1256 YFCISKRYLKEVRKPKYSMLGCMQFIPKDE 1287

RESULT 8
O9LAI3 PRELIMINARY; PRT; 451 AA.
AC O9LAI3;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Tetanus toxin (Fragment).
GN Clostridium tetani.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_Taxid=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-20886;
RA He H.J., Shi H.J., He Z.Y., Yuan Q.S., Wu X.F.;
RT "Fragment C of Tetanus Toxin.",
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154826; AAF73267.1; -.
DR HSSP; P04958; 1ABD.
DR InterPro: IPR001064; Crystallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 451 AA; 51823 MW; 69A8C5F0306C08E CRC64;

Query Match 22.3%; Score 162; DB 2; Length 451;
Best Local Similarity 27.5%; Pred. No. 6.6e-07;

```



```

Db      1258 CISKWYLKEVRRKRPYNSKLGCMWQETPRKDE 1287
          | ||:::| : |||::||| |
RESULT 12
09ZAU8 ID Q9ZAU8 PRELIMINARY; PRT; 1291 AA.
AC Q9ZAU8;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Bont protein.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1491;
  [1]
RN
RP SEQUENCE FROM N.A.
RA STRAIN=CDC 3281;
RX MEDLINE=98440323; PubMed=9767710;
RA Santos-Buelga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins."
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL; Y13630; CAA73968.1; -.
DR HSSP; P10845; 3BPA.
DR MEROPS; M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_Mprotease.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOKILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1291 AA; 150840 MW; E4D3B0E46AB2E735 CRC64;

Query Match 22.1%; Score 160.5; DB 2; Length 1291;
Best Local Similarity 28.0%; Pred. No. 3.3e-06;
Matches 42; Conservative 31; Mismatches 62; Indels 15; Gaps

QY 5 LYRGKFTIKKVASGN--KDNIVRNNDRYIVVWVKNKEKRLATNASQAGVEKILSATETI 62
      |||::|::: ||::| |::: ||::| |::|
Db 1139 LYIGKFTIIRKRSQSIINDIVREKDYILDYFNLNQMWRYYMKYFKKEBEKLFLEPI 1198
      ||:::|::: |:::| |::: |:::| |:::|

QY 63 PDVGYLMSOVVYVWVKSNDQCTINCKCM--NLQDNNGNDIGFIFGHPQNNIA-----KL 112
      : :|::| |::: |:::| |:::| |:::|
Db 1199 SDSDFEYNTIQTKEDEQP--TYSCLLFKKEESTDEIGLIHFRFYESGIVFKEYKDYF 1257
      | ||:::| : |||::||| |:::| |:::|

QY 113 VASNNYVNOIERS--SRTLGCSEFTIPVDD 140
      | ||:::| : |||::||| |:::| |:::|
Db 1258 CISKWYLKEVRRKRPYNSKLGCMWQETPRKDE 1287
      | ||:::| : |||::||| |

RESULT 13
093N27 ID Q93N27 PRELIMINARY; PRT; 1310 AA.
AC Q93N27;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Tetanus toxin (Fragment).
OS Clostridium tetani.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1513;
  [1]
RN
RP SEQUENCE FROM N.A.
RA Shumlin Z., Dianliang L.;
RT "Cloning and sequence analysis of tetanus toxin gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF389424; AAK72964.2; -.
DR InterPro: IPR000395; Bontoxilysin.

```

```

DR InterPro: IPR001064; CrystalLin.
DR InterPro: IPR000130; Zn_MTPEpse-.
DR Pfam: PF01742; Peptidase_M27.1.
DR ProDom: PD001963; BotoxILysin.1.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 1310 1310
SQ SEQUENCE 1310 AA; 150316 MW; 9EADDG914418E450 CRC64;

Query Match 22.0%; Score 160; DB 2; Length 1310;
Best Local Similarity 27.7%; Pred. No. 3.7e-06;
Matches 44; Conservative 24; Mismatches 49; Indels 42; Gaps 6;

QY 5 LYRGKFFIKKKYASGNK-DNIYRNDRYINVVYKREY-----RLA 45
D 1170 LYSGLKFTIKRTPNNELDSFVRSDPFIKLYSYNNHHIYGPCKDGNAFNNLDLIRYG 1229
QY 46 TNASOAGVEKLISALEIPDVGNLSOVVYKSKNDGCTNNCKMNIQDNNNGNDIGTIGH- 104
D 1230 YNAPGIPPIYKKMEAKVILKLDLK-----TYSVOLKLYLDDKASLGLVGTGN 1273
QY 105 -QFNNIK---LVASNNYRQIERSSRTLGCSWERTIPVD 139
D 1274 GQIGNDPNRDLIASNMYFNHLK--DKTLTCDMVFVPTD 1310

RESULT 14
ID 097FU1 PRELIMITARY; PRT: 315 AA.
AC 097FU1;
DT 01-OCT-2001 (TRENBLREL.18, Created)
DT 01-OCT-2001 (TRENBLREL.18, Last sequence update)
DT 01-OCT-2001 (TRENBLREL.18, Last annotation update)
DE Hypothetical protein CAC2635.
GN CAC2635.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gidson R., Lee H.M., Dubois J., Qiu D., Hilti J., Wolf Y.I.,
RA Tatusov R.L., Sabatne F., Doucetle-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007761; AAK80582.1;-
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 315 AA; 36592 MW; 1188CD3E83A22124 CRC64;

Query Match 15.1%; Score 110; DB 16; Length 315;
Best Local Similarity 32.6%; Pred. No. 0.028;
Matches 44; Conservative 21; Mismatches 52; Indels 18; Gaps 8;

QY 12 IIRKY---ASGNKDNIV--RNNDRYINVVYKREYRL-ATNASOAGVEKLISALE--I 62
D 47 IIRKYDYVGGDLKDNIVYIKINNNKSIIVKSKDKLYTLQPSKILNSLCTYKSYSPCLKIL 106
QY 63 PDVG--NLSQVYVYKSKKDDG-----ITNCKM-NLQDNNNGNDIGTGFPHQFNNAKLVA 114
D 107 TIDVSDNIDISLFTOSSENNMTNLQHLFIWDKTRFKDILHTNNIGFIDVHN-NKTPKITS 165
D 166 SNNFYRQIERSSRTL 129
D 166 SNNFYRQIERSSRTL 180

RESULT 15

```


O93HT3
 ID O93HT3 PRELIMINARY; PRT: 1291 AA.
 AC O93HT3.
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE Neurotoxin.
 GN NT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_taxid=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-YOICHI:
 RX MEDLINE-21534265; PubMed-11676492;
 RA Sagane Y., Kouguchi H., Watanabe T., Sunagawa H., Inoue K.,
 RA Fujinaga Y., Oguma K., Ohyama T.;
 RT "Role of C-Terminal Region of HA-33 Component of Botulinum Toxin in
 RT Hemagglutination.";
 RL Biochem. Biophys. Res. Commun. 288:650-657(2001).
 DR EMBL; AB061780; BAB71749.1; -
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_MTPeptide.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR ProDom; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN; 1.
 SQ SEQUENCE 1291 AA; 148869 MW; 4A21DB35B8743CF8 CRC64;

Query Match 13.8%; Score 100; DB 2; Length 1291;
 Best local similarity 23.6%; Pred. No. 1.3;
 Matches 37; Conservative 21; Mismatches 77; Indels 22; Gaps 4;

QY 2 NSSLYRGTKFIKKYASGNKDNIVANNDRVYINVVKKKEIRL-----ATNASQAGV 53
 Db 1139 NNDENEGYKIIIRKIRGNTINDIRVAGDILYFDMTINKKAYNLFMKNETMYADNHSTEDI 1198
 QY 54 EKLSALEIPVGNLSQVYVYKSKNDQGITNKC-KMNLQDNNGNDIGFIFGHQF----- 106
 Db 1199 YALGIREQTKINDNITIFQIQPMNNTYYASQIFKSNFNGENISGISIGYRFRLLGGDW 1258
 QY 107 --NNIAKLVASNMWYNRQIERSSRTLGCSEFIPYVD 140
 Db 1259 YRHNYLVPTVKQGNYSLSLETS---THMGFVPVSE 1291

Search completed: March 13, 2003, 11:40:13
 Job time : 23.4798 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:38:07 ; Search time 10.8934 Seconds
(without alignments)
1235.505 Million cell updates/sec

Title: US-09-917-791-22

Perfect score: 727

Sequence: 1 LNSSLYRGTFRILKKYASGN.....QIERSRFLGCSWEPIPYDD 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	100.0	1296	1	BRCCLAB
2	671	92.3	1296	2	botontoxilysin (EC 3
3	282.5	38.9	1268	2	botulinum neurotox
4	208.5	35.6	1274	2	botulinum neurotox
5	206.5	28.4	1252	2	botulinum neurotox
6	190	26.1	1251	2	botulinum neurotox
7	169.5	23.3	1291	2	botulinum neurotox
8	165.5	22.8	1291	1	botulinum neurotox
9	163	22.4	1315	1	botulinum neurotox
10	158.5	21.8	1297	2	botulinum neurotox
11	110	15.1	1315	2	botulinum neurotox
12	100	13.8	1291	2	botulinum neurotox
13	100	13.8	1291	2	botulinum neurotox
14	95	13.1	1285	2	botulinum neurotox
15	85.5	11.8	1285	2	botulinum neurotox
16	83	11.4	1285	2	botulinum neurotox
17	79.5	10.9	1285	2	botulinum neurotox
18	79	10.9	1285	2	botulinum neurotox
19	79	10.9	1285	2	botulinum neurotox
20	79	10.9	1285	2	botulinum neurotox
21	78.5	10.8	1285	2	botulinum neurotox
22	78.5	10.8	1285	2	botulinum neurotox
23	78.5	10.8	1285	2	botulinum neurotox
24	78.5	10.8	1285	2	botulinum neurotox
25	78	10.7	145	2	botulinum neurotox
26	78	10.7	145	2	botulinum neurotox
27	77.5	10.7	145	2	botulinum neurotox
28	77.5	10.7	145	2	botulinum neurotox
29	77	10.6	271	2	botulinum neurotox

30	77	10.6	4981	2	T18489	hypothetical prote
31	76.5	10.5	993	2	B64695	type I restriction
32	76.5	10.5	1278	2	B70236	hypothetical prote
33	76	10.5	472	2	AD2284	hypothetical prote
34	76	10.5	665	2	B71609	hypothetical prote
35	76	10.5	1175	2	F64489	hypothetical prote
36	75.5	10.4	233	2	AG1227	ABC transporter, A
37	75.5	10.4	323	2	C83940	sugar ABC transpor
38	75	10.3	257	2	A96908	ABC transporter, A
39	75	10.3	708	2	A95214	cell wall surface
40	75	10.3	719	2	B98078	hypothetical prote
41	75	10.3	732	2	C64972	probable ATPase -
42	75	10.3	1247	2	E71616	hypothetical prote
43	75	10.3	1711	2	T18429	hypothetical prote
44	74.5	10.2	233	2	C95192	ABC transporter, A
45	74.5	10.2	435	2	A11017	probable type-I se

ALIGNMENTS

RESULT 1
BRCCLAB
bontoxilysin (EC 3.4.24.69) - A precursor - Clostridium botulinum
N:Alternate names: botulinum neurotoxin type A
C:Species: Clostridium botulinum
C:Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 18-Jun-1999
C:Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R:Blitz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A:Reference number: A35294; MUID:90264400; PMID:2160960
A:Accession: A35294
A:Molecule type: DNA
A:Residues: 1-1296 <BIN>
A:Cross-references: GB:M30196; NID:9144864; PIDN:AAA23262.1; PID:9144865
A:Experimental source: strain 62A, subtype A
R:Thompson, D.E.; Brem, J.K.; Oultam, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neuroto
A:Reference number: S09492; MUID:90235864; PMID:2185020
A:Accession: S09492
A:Molecule type: DNA
A:Residues: 1, 'Q', 3-26, 'V', 28-1296 <THO>
A:Cross-references: EMBL:X52066; NID:940381; PIDN:CAA36289.1; PID:940382
A:Experimental source: NCTC 2916
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone
A:Reference number: S67988; MUID:96096783; PMID:8521962
A:Accession: S67988
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <FUJ>
A:Cross-references: EMBL:D67030; DBJ:D50421; NID:92160224
A:Experimental source: B.R.
R:Betley, M.J.; Somers, E.; Dasgupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-t
A:Reference number: A33401; MUID:89350959; PMID:2669749
A:Accession: A33401
A:Molecule type: DNA
A:Residues: 1-35 <BET>
A:Cross-references: GB:M27892; NID:9144880; PIDN:AAA23269.1; PID:9551776
R:Gimenez, J.A.; Dasgupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
A:Reference number: A53884; MUID:94000342; PMID:8397793
A:Accession: A53884
A:Status: preliminary
A:Molecule type: protein
A:Residues: 867-880;1148-1217, 'Y', 1219 <GIN>
A:Experimental source: strain Hall
A:Note: sequence extracted from NCBI backbone (NCBI:139159); sequence modified after

R:Das Gupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and around
A:Reference number: A60025; MUID:91120847; PMID:2126206
A:Accession: A60025
A:Molecule type: protein
A:Residues: 2-6:445-453, 'X', 455-457 <DAS1>
R:Das Gupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A:Reference number: A27000
A:Accession: A27000
A:Molecule type: protein
A:Residues: 2-47 <DAS2>
R:Binz, T.; Blaszi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
A:Reference number: A49708; MUID:94124495; PMID:8294407
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synapses
C:Genetics:
A:Gene: atx; botA
A:Function:
C:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associated
C:Superfamily: tetanus toxin
C:Keywords: disulfide bond; hydrolase; metalloprotease; neurotoxin; transmembrane protein
F:2-444/Product: botoxilysin A light chain #status experimental <LGHT>
F:445-1296/Product: botoxilysin A heavy chain #status experimental <HVT>
F:223,227/Binding site: zinc (His) #status predicted
F:224/Active site: Glu #status predicted

Query Match 100.0%; Score 727; DB 1; Length 1296;
Best Local Similarity 100.0%; Pred. No. 5.9e-58;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNSSLKRGTFIKKAYASGKNDIVRNDRVYINVVVKKKEYRLATNASQAGVEKILTSAL 60
|||||
DB 1150 LNSSLKRGTFIKKAYASGKNDIVRNDRVYINVVVKKKEYRLATNASQAGVEKILTSAL 1209

QY 61 EIPDVGNLSQVVMKSKNDGKITNCKKMLQDNGNDIGFIFGHQFNNTAKLVAASWYNR 120
|||||
DB 1210 EIPDVGNLSQVVMKSKNDGKITNCKKMLQDNGNDIGFIFGHQFNNTAKLVAASWYNR 1269

QY 121 QIERSSRTLCGSWEFTIPVD 140
|||||
DB 1270 QIERSSRTLCGSWEFTIPVD 1289

RESULT 2
140645
botulinum neurotoxin type A - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40645
R:Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A:Reference number: I40645; MUID:94143603; PMID:8310180
A:Accession: I40645
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1296 <RES>
A:Cross-References: EMBL:X73423; NID:g507070; PIDN:CAA51824.1; PID:g507071
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 92.3%; Score 671; DB 2; Length 1296;
Best Local Similarity 90.7%; Pred. No. 7.3e-53;
Matches 127; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 LNSSLKRGTFIKKAYASGKNDIVRNDRVYINVVVKKKEYRLATNASQAGVEKILTSAL 60
|||||
DB 1150 LNSSLKRGTFIKKAYASGKNDIVRNDRVYINVVVKKKEYRLATNASQAGVEKILTSAL 1209

QY 61 EIPDVGNLSQVVMKSKNDGKITNCKKMLQDNGNDIGFIFGHQFNNTAKLVAASWYNR 120
|||||
DB 1210 EIPDVGNLSQVVMKSKNDGKITNCKKMLQDNGNDIGFIFGHQFNNTAKLVAASWYNR 1269

QY 121 QIERSSRTLCGSWEFTIPVD 140
|||||
DB 1270 QVCKASRTFCGSWEFTIPVD 1289

RESULT 3
S33411
botulinum neurotoxin type F - Clostridium baratii
C:Species: Clostridium baratii
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411; S31860
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.
FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin
A:Reference number: S33411; MUID:93252228; PMID:8486245
A:Accession: S33411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-References: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 38.9%; Score 282.5; DB 2; Length 1268;
Best Local Similarity 45.3%; Pred. No. 1.5e-17;
Matches 63; Conservative 19; Mismatches 44; Indels 13; Gaps 5;

QY 2 NSSLYGTFIKKAYASGKNDIVRNDRVYINVVVKKKEYRLATNASQAGVEKILTSAL 58
|||||
DB 1131 NALVYGVETILKVGSTSTSDNDFVRKNDYINVVGVGNSFYLYAVTSASVAKITK 1190

QY 59 ALEIPVGNLSQVVMKSKNDGKITNCKKMLQDNGNDIGFIFGHQFNNTAKLVAASWYNR 117
|||||
DB 1191 LRRISNSNYNSQMIIMDS----IGDCTMNFKTNNGNDIGLGFH-LNN--LVAASW 1241

QY 118 YNRQIERSSRTLCGSWEFTI 136
|||||
DB 1242 YKNIIRNTRNNGCFWFSFI 1260

RESULT 4
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, P.
FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40813
A:Accession: I40813
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-References: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48108
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAM>
A:Cross-References: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 35.6%; Score 258.5; DB 2; Length 1274;

Best Local Similarity 43.2%; Pred. No. 2.4e-15;
Matches 60; Conservative 20; Mismatches 44; Indels 15; Gaps 5;

Qy 1 LNSSLYRGTKFLIKRYAS---GNKNDIVRNNDRYIINVVYVVKNEKRLATNMQAGVEKIL 57
Db 1140 LNKLYKSEVEVYIRKNGPIDISNTDNEVFRKNDLAINVDRGVEYRLAD--TKSEKRII 1198
Qy 58 SALEIPDVGNLSQVYVVMKSKNDGKITNCKMNLQDNGNDIGFIFGHOFPNNIAKLVASNW 117
Db 1199 RTSNND--SLGQITVMS-----IGNNCTMNFQNNKNSNIGLGFHSNN-----LVASSW 1247
Qy 118 YNRQIERSRRLGCSWEFI 136
Db 1248 YNNIRRTSSNGCFWSSI 1266

RESULT 5

S21178
C:Species: Clostridium botulinum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111
R:Wheeler, S.M.; Elmore, M.J.; Bodsworth, N.D.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992
A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
A:Reference number: S21178; MUID:92174922; PMID:1541280
A:Accession: S21178
A:Molecule type: DNA
A:Residues: 1-1252 <MHE>
A:Cross-references: EMBL:X62683; NID:940397; PIDN:CAA44558.1; PID:940398
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48107
A:Molecule type: DNA
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 616-982 <CAM>
A:Cross-references: EMBL:X70815; NID:9407786; PIDN:CAA50146.1; PID:9407787
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
A:Reference number: JH0256; MUID:92181428; PMID:1543481
A:Accession: JH0257
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'I', 774-962, 'FE', 965-966, 'R', 968-1
A:Cross-references: EMBL:X62089; NID:940393; PIDN:CAA43999.1; PID:940394
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A:Reference number: A35294; MUID:90264400; PMID:2160960
A:Accession: B35294
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-252 <BIN>
A:Experimental source: strain Beluga
R:Gienez, J.A.; Dasgupta, B.R.
Biochimie 72, 213-217, 1990
A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the s
A:Reference number: A60027; MUID:90344918; PMID:2116911
A:Accession: A60027
A:Molecule type: protein
A:Residues: 420-427 <GIM>
A:Experimental source: strain Beluga
A:Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin
C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>

F:423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>
F:412-426/Dissulfide bonds: #status predicted

Query Match 28.4%; Score 206.5; DB 2; Length 1251;
Best Local Similarity 38.1%; Pred. No. 1.2e-10;
Matches 53; Conservative 20; Mismatches 51; Indels 15; Gaps 6;
Qy 1 LNSSLYRGTKFLIKRY-ASGNKNDIVRNNDRYIINVVYVVKNEKRLATNMQAGVEKILS 58
Db 1117 LANRLYSIGIKYIQVNNSSSTNDNLVVRKNDQYIINFVASKHLPFLYADTATNKEK--- 1173
Qy 59 ALEIPDVGNLSQVYVVMKSKNDGKITNCKMNLQDNGNDIGFIFGHOFPNNIAKLVASNW 117
Db 1174 TIKISSGNRFNQVYVVMNS-----VGNCTMNFQNNKNNNGNIGLGF-----KADTVASTW 1224
Qy 118 YNRQIERSRRLGCSWEFI 136
Db 1225 YTHMRDHTNSNGCFWNPFI 1243

RESULT 6

JH0256
C:Species: Clostridium butyricum
C:Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
C:Accession: JH0256; S16145
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum t
A:Reference number: JH0256; MUID:92181428; PMID:1543481
A:Accession: JH0256
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-27, 'E', 29-1251 <POU>
A:Cross-references: EMBL:X62088; NID:940379
A:Experimental source: strains ATCC 43181 and ATCC 43755
R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa,
J. Gen. Microbiol. 137, 519-525, 1991
A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E t
A:Reference number: S16145; MUID:91237316; PMID:2033376
A:Accession: S16145
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229, 'M', 231-252 <FNU>
A:Cross-references: EMBL:X53180; NID:940407; PIDN:CAA37321.1; PID:940408
A:Experimental source: strain BL6340
C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter relea
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
F:423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>
F:412-426/Dissulfide bonds: #status predicted

Query Match 26.1%; Score 190; DB 2; Length 1251;
Best Local Similarity 37.7%; Pred. No. 3.9e-09;
Matches 52; Conservative 19; Mismatches 53; Indels 14; Gaps 6;

Qy 1 LNSSLYRGTKFLIKRY-ASGNKNDIVRNNDRYIINVVYVVKNEKRLATNMQAGVEKILS 59
Db 1117 LANRLYSIGIKYIQVNNSSSTNDNLVVRKNDQYIINFVA-SKTHLLPLVADTATNKEK-EKT 1174
Qy 60 LEIPDVGNLSQVYVVMKSKNDGKITNCKMNLQDNGNDIGFIFGHOFPNNIAKLVASNWY 118
Db 1175 IKISSGNRFNQVYVVMNS-----VGNCTMNFQNNKNNNGNIGLGF-----KADTVASTWY 1224
Qy 119 NROJERSRRLGCSWEFI 136
Db 1225 YTHMRDHTNSNGCFWNPFI 1242
RESULT 7
140631

[illegible]

A:Experimental source: type B, Danish
A>Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBIF:112081); this publ
R:Campbell, K.D.; Collins, M.D.; Best, A.K.
J: Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372; PMID:8408542
A:Accession: S48105
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 634-994 <CAN>
A:Cross-references: EMBL:X70817; NID:9407782; PIDN:CAAS0148.1; PID:9407783
A:Experimental source: proteolytic type B, strain NCTC 7273
R:Szabo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
submitted to the EMBL Data Library, April 1992
A:Description: Partial amino acid sequence of botulinum neurotoxin type B and compari
A:Reference number: S21575
A:Accession: S21575
A:Molecule type: DNA
A:Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZA>
A:Cross-references: EMBL:Z11934; NID:940383; PIDN:CAAT7991.1; PID:940384
R:Kuzanov, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars
J: Biol. Chem. 267, 14721-14729, 1992
A:Title: Minimal essential domains specifying toxicity of the light chains of tetanus
A:Reference number: A42871; MUID:92340509; PMID:1654516
A:Accession: A42871
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-313, 'S', 315-451 <KUR>
A:Experimental source: strain Okra
A>Note: sequence extracted from NCBI backbone (NCBIF:109365)
R:Dasgupta, B.R.; Datta, A.
Biochimie 70, 811-817, 1988
A:Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity wi
A:Reference number: S07155; MUID:89000987; PMID:3139097
A:Accession: S07155
A:Molecule type: protein
A:Residues: 2-29, 'W', 31-45 <DAS>
A:Accession: S08562
A:Molecule type: protein
A:Residues: 442-463, 'R', 465-467 <DA2>
R:Schmidt, U.J.; Satyamorchy, V.; Dasgupta, B.R.
Arch. Biochem. Biophys. 238, 344-348, 1985
A:Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A:Reference number: S07128; MUID:85197963; PMID:388113
A:Accession: S07128
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-16 <SCH1>
A:Accession: S08573
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-17 <SCH2>
A:Accession: S08574
A:Status: preliminary
A:Molecule type: protein
A:Residues: 442-459 <SCH3>
R:Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,
Nature 359, 832-835, 1992
A:Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MUID:93065293; PMID:1331807
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C:Genetics:
A:Gene: hont/b
C:Function:
A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C:Superfamily: tetanus toxin
C:Keywords: hydrolysis; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-441/Product: bontoxilysin B light chain #status experimental <GNT>
F:442-1291/Product: bontoxilysin B heavy chain #status experimental <HNT>
F:230, 234/Binding site: zinc (HIS) #status predicted
F:231/Active site: Gln #status predicted

```

Query Match      22.8%   Score 165.5; DB 1; Length 1291;
Best Local Similarity 28.7%; Pred. No. 6.9e-07;
Matches 43; Conservative 31; Mismatches 61; Indels 15; Gaps 5;

OY 5 LVGKTFIIKKVSGN--KDNIVRNNDREVIINVVYKNENKEEYRLATNASQGVETLTALEI 62
    ||| |||| : : : ||| | : : : | : | : | : | : | : | : | : |
DB 1139 LVIGKEFTIRKSNSGSINDIVAKEDYITIDFNLDMWRKYTTTKYFKKEEKLEFLAPI 1198

OY 63 PDVGNLSQVVVMKSKNDOGITNKCM--NIODNNNGDIGNCIEFGHFNNTA-----KL 112
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1199 SDSEDFNYNTIQIKVEYDQP-TYSCQLLFKKDESESTDELGLIHRYESGIYEYKDYF 1257

OY 113 VASNMWNRQIER--SRTLGCSEFIPIVD 140
    || : : : : : || : : : : : || : : : : : || : : : : : || :
DB 1258 CISKWYLKEVKRRPYNLKLCNMQIFPKDE 1287

RESULT 9
BICLRN
tentoxilysin (EC 3.4.24.68) precursor - Clostridium tetani
N.Alternate names: tetanus neurotoxin
C.Species: Clostridium tetani
C.Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-Jun-2002
C.Accession: A25689, A25757, A25194, B25194, A60759, S69348, S09364
R.Eisel, U.; Jarausch, W.; Goretzki, K.; Henschen, A.; Engels, J.; Weller, U.; Hudel, M.
EMBO J. 5, 2495-2502, 1986
A.Title: Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxin
A.Reference number: A25689; MWID:87055814; PMID:3536478
A.Accession: A25689
A.Molecule type: DNA
A.Residues: 1-1315 <EAIS>
A.Cross-references: GB:X04436; NID:g40769; PIDN:CAA28033.1; PID:g40770
R.Fairweather, N.F.; Lyness, V.A.
Nucleic Acids Res. 14, 7809-7812, 1986
A>Title: The complete nucleotide sequence of tetanus toxin.
A.Reference number: A25757; MWID:87040747; PMID:3774547
A.Accession: A25757
A.Molecule type: DNA
A.Residues: 1-1315 <FAI>
A.Cross-references: GB:X06214; NID:g40773; PIDN:CAA29564.1; PID:g40774
A.Experimental source: strain CN3911
R.Fairweather, N.F.; Lyness, V.A.; Pickard, D.J.; Allen, G.; Thomson, R.O.
J. Bacteriol. 165, 21-27, 1986
A>Title: Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in Escherichia coli.
A.Reference number: A25194; MWID:86085672; PMID:3510187
A.Accession: A25194
A.Molecule type: DNA
A.Residues: 743-1315 <FAZ>
A.Cross-references: GB:M12739; NID:g144920; PIDN:AAA23282.1; PID:g144921
A.Accession: B25194
A.Molecule type: protein
A.Residues: 865-894 <FAZ>
R.Matsuda, M.; Iel, D.L.; Sugimoto, N.; Ozutsumi, K.; Okabe, T.
Infect. Immun. 57, 3588-3593, 1989
A>Title: Isolation, purification, and characterization of fragment B, the NH-2-terminal portion of tetanus toxin.
A.Reference number: A60759; MWID:90035436; PMID:2478476
A.Accession: A60759
A.Molecule type: protein
A.Residues: 461-475 <MAT>
R.Demotz, S.; Lanzavecchia, L.; Eisel, U.; Niemann, H.; Widmann, C.; Corradin, G.
J. Immunol. 142, 394-402, 1989
A>Title: Determination of several DR-restricted tetanus toxin T cell epitopes.
A.Reference number: JS0098; MWID:89093918; PMID:2463305
A.Contents: annotation; epitope region
R.Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, B.R.;
Eur. J. Biochem. 229, 61-69, 1995
A>Title: Structural studies on the zinc-endopeptidase light chain of tetanus neurotoxin.
A.Reference number: S69348; MWID:95262688; PMID:7744050

```

[illegible]

Query Match	15.1%	Score 110;	DB 2;	Length 315;
Best Local Similarity	32.6%;	Pred. No. 0.014;		
Matches 44; Conservative	21;	Mismatches 52;	Indels 18;	Gaps 8

[illegible]

RESULT 12
 S46431
 botulinum neurotoxin C1 - Clostridium botulinum phage 1C (strain C 468)
 N:Alternate names: BoNT/C1 protein
 C:Species: Clostridium botulinum phage 1C
 A:Variety: strain C 468
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
 C:Accession: S46431; S49107
 R:Hauser, D.; Eklund, M.W.; Boquet, P.; Popoff, M.R.
 M1. Gen. Genet. 243, 631-640, 1994
 A:Title: Organization of the botulinum neurotoxin C1 gene and its associated non-toxic P
 A:Reference number: S46426; MUID:94301293; PMID:8028579
 A:Accession: S46431
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1291 <HAUV>
 A:Cross-References: EMBL:X72793; NID:g516171; PIDN:CAAS1313.1; PID:g516175
 A:Experimental source: strain C 468
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
 C:Superfamily: tetanus toxin

Query Match	13.8%	Score 100:	DB 2;	Length 1291;
Best Local Similarity	23.6%	Pred. No.	0.62;	
Matches 37; Conservative	21;	Mismatches	77;	Indels 22; Gaps 4

```

0Y 2 NSSLYRGKFLIKKAYASGNKDNITRNDRYINVVKKNEYRL-----ATMNASQAV 53
Db 1139 NNDENEGRIKIKRIKGNJTMDTRRGGDILLYPMITINNKAYILPMKNETMYADNHNSTEDI 1196
0Y 54 EKLSALEIPDVGLSQVYVKKSKNDGKITNKC-KNMLDNNKNDIGFTGPHQF----- 106
Db 1199 YALGIREFTKIDINIIFFQIQPMNNIYYASQIFKSNFGENISIGISIGTGYFRILGDW 1256
0Y 107 ---NNIAKLVASNNYRNQIERSSRTLCGSEFIYPVD 140
Db 1259 YRHNTLVPTVAKGNVASTLEST---THNGEYPAVSE 1291

```

RESULT 13
A49777

C:Species: Clostridium botulinum phage
C:Date: 10-Mar-1994 #sequence_revision 07-Apr-1994 #text_change 23-Mar-2001
C:Accession: S11291, A35396, S22166, M49777
R:Hauser, D., Eklund, M.W., Kurazono, H., Binz, T., Niemann, H., Gill, D.M., Boquet, L.
Nucleic Acids Res. 18, 4924, 1990
A:Title: Nucleotide sequence of Clostridium botulinum C1 neurotoxin.
A:Reference number: S11291; MUID:90370487; PMID:2204031

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84, 'P', 86-1291 <H4U>
A:Cross-references: EMBL:X53751; NID:g14905; PIDN:CAA37780.1; PID:g14906
R:Kimura, K.; Fujii, N.; Tsuchi, K.; Murakami, T.; Indoh, T.; Yokosawa, N.; Takeshi,
Biochem. Biophys. Res. Commun. 171, 1304-1311, 1990
A:Title: The complete nucleotide sequence of the gene coding for botulinum type C-1 toxin
A:Reference number: A35396; MUID:9102498; PMID:2222445
A:Accession: A35396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA

R;TsuZuki, K.; Kimura, K.; Fujii, N.; Yokosawa, N.; Oguma, K.
submitted to the EMBL Data Library, December 1991
A:Description: Nucleotide sequence of the gene for one of the components of hemagglut
A:Reference number: S22163

A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-1291 <TS2>
A: Cross-references: EMBL:X63389; NID:9558175; PID:940390
R: Kimura, K., Fujii, N., Tsuchizaki, K., Murakami, T., Indo, T., Yokosawa, N., Oguma, K.
Appl. Environ. Microbiol. 57, 1168-1172, 1991
A: Title: Cloning of the structural gene for Clostridium botulinum type C-1 toxin and
Reference number: A49777; MUID:91282468; PMID:2059039

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-607 <TS3>
A;Cross-references: GB:D90210
C;Superfamily: tetanus toxin
C;Keywords: neurotoxin

Query Match	13.8%	Score 100;	DB 2;	Length 1291;
Best Local Similarity	23.6%	Pred No. 0.62;		
Matches 37; Conservative	21;	Mismatches	77;	Indels 22; Gaps 4

```
Qy      2 NSSLRGCTFLIKKAYAGNNKDIVRNNDRYINVVYKKNEYL-----ATNMQAGV 53
          | : | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db     1139 NNDENEGKIIIKIRGNTDTRRCGGDLIFEMTITNNKAYILFKMETMYADHNSTEDI 1198
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy      54 EKILSALEIPVGMLSQVAVMKSKNDQGITNKC-KKNLQDNNGNDIGFGRHO----- 106
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     1199 YATLRGTQKOINDNIITIQAPMNNYYIASQLFKSNENGENISGICSTGYRFRLGDM 1258
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy      107 ---NNIAKLVAAMNYNQIERSSRTLGCSWEFIYPD 140
          | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     1259 YRHANYLVPTVKQGNVASLESTS----THKGEPVASE 1291
```

RESULT 14
S70582
botulinum neurotoxin type Dsa precursor - Clostridium botulinum phage d-sa
C:Species: Clostridium botulinum phage d-sa
A:Note: host Clostridium botulinum type D (strain South Africa)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000
C:Accession: S70582
R:Morishih, K.; Koura, M.; Abe, N.; Fujii, N.; Fujinaga, Y.; Inoue, K.; Ogumad, K.
Biochim. Biophys. Acta 1307, 123-126, 1996
A:Title: Mosaic structures of neurotoxins produced from Clostridium botulinum types C
;Reference number: S70582; MUID:96283801; PMID:8679691

A:Accession: S70582
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1285 <MOR>
A:Cross-references: EMBL:D38442; NID:q1374775; PIDN:BA07477.1; PID:q1374776
C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
a disulfide bond. The heavy chain mediates the binding of toxin to the presynaptic membr
C:Superfamily: tetanus toxin
C:Keywords: disulfide bond; neurotoxin; transmembrane protein
F:1-447/Product: botulinum neurotoxin type Dsa light chain #status predicted <MAT1>
F:448-1285/Product: botulinum neurotoxin type Dsa heavy chain #status predicted <MAT2>

Query Match 13.1%; Score 95; DB 2; Length 1285;

Best Local Similarity 23.8%; Pred. No. 1.8;
Matches 38; Conservative 20; Mismatches 72; Indels 30; Gaps 7;

OY 2 NSSLYRGTKFIKKYASGKNDIVNNDRYINVVYKNEYRLATNASQAGYEKILSALE 61

Db 1135 NNDFFEGYKIIIRKRGNTNDTRVGENVLYENTYIDNKQYSLGMYKPSRNL----GTD 1189

OY 62 IPDVGNLSQVYVYVMSKNDGITNKC-----KMNLDN-NGNDIGF--IGFHQFNNTA 110

Db 1190 LVPDLALDQPMDEIRKYSFTIQPCNTDYASQLPLSSNATYTRKGLISTGSYSF---- 1245

OY 111 KLVASNMWYNRQ-----IERSSRYL--GCSWEFIPYDD 140

Db 1246 KLGDDYWFNHEYLIPVIEHYASLSLESTHWFVPASE 1285

RESULT 15

D96901

probable dehydrogenase with iron-sulfur domain [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: D96901

R:Noiling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium CLO

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D96901

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-467 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK77999.1; PID:q15022830; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0012

Query Match 11.8%; Score 85.5; DB 2; Length 467;

Best Local Similarity 28.8%; Pred. No. 3.9;
Matches 40; Conservative 25; Mismatches 51; Indels 23; Gaps 8;

OY 1 INSSLYRGTKFIKKYASGKNDIVNNDRYINVVYK-NKEYRLATNASQAGVEKILSA 59

Db 26 INVSL-----IEKEYEIGDPTSIY-NSDLVYDGIESKYNFNYGLETEGNOE-IEKICSV 77

OY 60 LEIPVGNLSQVYVYVMSKNDGICITNK-----CKMNLQDNGNDIGFISGHQFNNTAK-LV 113

Db 78 LSVF-FKRVNMAIVCSRDENEEDTINKIYSGIKRGIKIDASLANIE-----SAIAKSIIT 129

OY 114 ASNMWYNRQIERSSRYLGS 132

Db 130 LKDIINNKIELEKDIGIVCA 148

Search completed: March 13, 2003, 11:41:46
Job time : 12.8934 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 11:36:37 ; Search time 9.54467 Seconds
(without alignments)
899.518 Million cell updates/sec

Title: US-09-917-791-21

Perfect score: 1071

Sequence: 1 IKVNNMDLFFSPSEDNFTND.....NIGNNMLYKDFVGLIFSGA 207

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1071	100.0	1295	1 BXA1_CLOBO	P10845 clostridium
2	918	85.7	1295	1 BXA2_CLOBO	Q45894 clostridium
3	377	35.2	1296	1 BXG_CLOBO	O60393 clostridium
4	374.5	35.0	1250	1 BXE_CLOBO	P30995 clostridium
5	368.5	34.2	1250	1 BXE_CLOBO	O00496 clostridium
6	366.5	34.2	1290	1 BXA_CLOBO	P10844 clostridium
7	316	29.5	1274	1 BXA_CLOBO	P30996 clostridium
8	312	29.1	1314	1 TFX_CLOBO	P04958 clostridium
9	290.5	27.1	1276	1 BXD_CLOBO	P19321 clostridium
10	263	24.6	1290	1 BXD_CLOBO	P16640 clostridium
11	107	10.0	246	1 Y402_BUCAI	P57482 buchnera ap
12	106	9.9	1196	1 BXC1_CLOBO	P46081 clostridium
13	98	9.2	1345	1 BXC1_CLOBO	P33366 saccharomyc
14	95.5	8.9	1032	1 Y198_YEAST	P40469 saccharomyc
15	93.5	8.7	355	1 Y198_YEAST	O9467 rickettsia
16	91	8.5	635	1 HTPG_VIBCH	P22359 vibrio chol
17	91	8.5	1162	1 BXA_CLOBO	O06366 clostridium
18	90.5	8.5	1186	1 BXA_CLOBO	P55980 helicobacte
19	89	8.3	1162	1 BXA_CLOBO	P46082 clostridium
20	88.5	8.3	556	1 BXA_CLOBO	O93712 listeria in
21	87	8.1	755	1 KHL5_HUMAN	O96P97 homo sapien
22	86.5	8.1	556	1 SYR_LISMO	O8Y493 listeria mo
23	86.5	8.1	1939	1 MYH4_HUMAN	O9Y623 homo sapien
24	86	8.0	610	1 MYH4_HUMAN	O51229 botreilla bu
25	86	8.0	844	1 SECA_STACA	P47994 staphylococ
26	86	8.0	891	1 SECA_PAVLU	O01570 pavlova lut
27	86	8.0	974	1 TRP4_MOUSE	O9495 mus musculu
28	86	8.0	977	1 TRP4_MOUSE	O9494 mus musculu
29	86	8.0	977	1 TRP4_HUMAN	O35119 rattus norv
30	86	8.0	981	1 TRP4_BOVIN	P79100 bos taurus
31	85.5	8.0	426	1 HEM1_SULSO	O960U7 sulfolobus
32	85.5	8.0	1167	1 HEM1_SULSO	O921T1 helicobacte
33	85	7.9	336	1 Y05G_BPT4	P39242 bacteriophage

ALIGNMENTS

RESULT 1	ID	Sequence	Standard	PRT	AA
AC	BXA1_CLOBO	P10845: P18639; P01561;			
DT	01-JUL-1988 (Rel. 11, created)				
DT	01-JUL-1993 (Rel. 26, last sequence update)				
DT	15-JUN-2002 (Rel. 41, last annotation update)				
DE	Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)				
DE	chain; Botulinum neurotoxin A, heavy-chain.				
GN	BOTA OR BNA OR AFX				
OS	Clostridium botulinum				
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.				
OX	NCBI_TaxID=1491;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NCCT 2916;				
RX	MEDLINE=90235864; PubMed=2185020;				
RA	Thompson D.E., Brehm J.K., Oultram J.D., Swinfield T.J.,				
RA	Shone G.C., Atkinson T., Melling J., Minton N.P.;				
RT	"The complete amino acid sequence of the Clostridium botulinum type A neurotoxin, deduced by nucleotide sequence analysis of the encoding gene.";				
RT	Eur. J. Biochem. 189:73-81(1990).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=62A.				
RX	MEDLINE=90264400; PubMed=2160960;				
RA	Binz B., Kuatzen H., Wille M., Frevent J., Wernars K., Niemann H.;				
RT	"The complete sequence of botulinum neurotoxin type A and comparison with other clostridial neurotoxins.";				
RT	J. Biol. Chem. 265:9153-9158(1990).				
RL	[3]				
RP	SEQUENCE OF 1-65 FROM N.A.				
RC	STRAIN=62A.				
RX	MEDLINE=97016817; PubMed=8863443;				
RA	Fast A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;				
RT	"Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component.";				
RT	Int. J. Syst. Bacteriol. 46:1105-1112(1996).				
RL	[4]				
RP	SEQUENCE OF 1-34 FROM N.A.				
RC	STRAIN=Hall.				
RX	MEDLINE=89350959; PubMed=2669749;				
RA	Betley M.J., Somers E., Dasgupta B.R.;				
RT	"Characterization of botulinum type A neurotoxin gene: delineation of the N-terminal encoding region.";				
RT	Biochem. Biophys. Res. Commun. 162:1388-1395(1989).				
RL	[5]				
RP	SEQUENCE OF 1-18 FROM N.A.				
RC	STRAIN=Type A NIH;				
RX	MEDLINE=96096783; PubMed=8521962;				
RA	Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;				

"Molecular characterization of two forms of nontoxic-nonhemagglutinin components of Clostridium botulinum type A progenitor toxins.";
FEBS Lett. 376:41-44(1995).
[6]
RP SEQUENCE OF 1-16.
RX MEDLINE-84178501; PubMed-6370252;
RA Schmidt J.J., Sarcymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequence of the heavy and light chains of botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RA Dasgupta B.R., Foley J., Niece R.;
RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE-91120847; PubMed-2126206;
RA Dasgupta B.R., Dekleva M.L.;
RT "Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and around the nicking site.";
RL Biochimie 72:661-664(1990).
RN [9]
RP SEQUENCE OF 448-464 AND 872-895.
RX MEDLINE-89024662; PubMed-3178218;
RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;
RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two halves and their partial sequences.";
RL Arch. Biochem. Biophys. 266:142-151(1988).
RN [10]
RP SEQUENCE OF 448-482.
RX MEDLINE-85285016; PubMed-3896784;
RA Shone C.C., Hambleton P., Melling J.;
RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity.";
RL Eur. J. Biochem. 151:75-82(1985).
RN [11]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE-94063091; PubMed-8243676;
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J., Benfenati F., Wilson M.C., Montecucco C.;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN [12]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE-94124495; PubMed-8294407;
RA Jahn R., Biasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
RN [13]
RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX MEDLINE-21556941; PubMed-11700044;
RA Rigoni M., Caccini P., Johnson E.A., Montecucco C., Rossetto O.;
RT "Site-directed mutagenesis identifies active-site residues of the light chain of botulinum neurotoxin type a.";
RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE-98455071; PubMed-9783750;
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT "Crystal structure of botulinum neurotoxin type A and implications for toxicity.";
RL Nat. Struct. Biol. 5:898-902(1998).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin binds with high affinity to peripheral neuronal presynaptic membrane, is then internalized by receptor-mediated endocytosis. The C-terminus of the heavy chain (H) is responsible for the adherence of the toxin to the cell surface while the N-terminus mediates transport of the light chain from the endocytic vesicle

to the cytosol. After translocation, the light chain (L) hydrolyzes the 197-Gln-I-Arg-198 bond in SNAP-25, thereby blocking neurotransmitter release. Inhibition of acetylcholine release results in flaccid paralysis, with frequent heart or respiratory failure.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PHARMACEUTICAL: Available under the name BOTOX (Allergan) for the treatment of strabismus and blepharospasm associated with dystonia and cervical dystonia. Also used for the treatment of hemifacial spasm and a number of other neurological disorders characterized by abnormal muscle contraction.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -1- DATABASE: NAME-BOTOX product information Web site; WWW="http://www.botox.com/index.jsp?hp&productinfo".
CC -1- DATABASE: NAME-Protein Spotlight;
CC NOTE-Issue 19 of February 2002;
CC WWW="http://www.expasy.org/spotlight/articles/spl1019.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52066; CAA36289.1; -
DR EMBL: M30196; AAA3262.1; -
DR EMBL: X82973; CAA63551.1; -
DR EMBL: D67030; BAA11051.1; -
DR EMBL: M27892; AAA3269.1; -
DR PIR: A35294; B7CLAB.
DR PIR: S09492; S09492.
DR PIR: 3BTA; 01-OCT-99.
DR MEROPS: M27.002; -
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptidse.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW Pharmacological; 3D-structure.
FT INIT_MET 0
FT CHAIN 1 447
FT METAL 448 1295
FT ACT_SITE 222 222
FT METAL 222 222
FT ACT_SITE 223 223
FT METAL 226 226
FT METAL 261 261
FT DISULFID 429 453
FT DISULFID 1234 1279
FT TRANSMEM 626 646
FT TRANSMEM 655 675
FT VARIANT 26 26
FT VARIANT 261 261
FT MUTAGEN 265 265
FT MUTAGEN 365 365
FT CONFLICT 1 1
FT CONFLICT 479 479
FT CONFLICT 875 875
FT CONFLICT 891 891
SQ SEQUENCE 1295 AA; 149322 MW; 858342F754862579 CRC64;
Query Match 100.0%; Score 1071; DB 1; Length 1295;
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).
INTERCHAIN.
POTENTIAL.
POTENTIAL.
V->A.
E->A: DRASTIC DECREASE IN ENZYMATIC ACTIVITY.
F->A: DECREASE IN ENZYMATIC ACTIVITY.
Y->A: DECREASE IN ENZYMATIC ACTIVITY.
P->Q (IN REF. 1).
E->P (IN REF. 9).
T->L (IN REF. 8).
S->K (IN REF. 8).
ZINC (CATALYTIC).
ZINC (CATALYTIC).<

Best Local Similarity 100.0%; Pred. No. 1.5e-78;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 IKVNMNDLFFSPSEDNTNDLNGEETSDNTNEAEENISLDLQOYLTFFNDPEN 60
CC |
CC 454 IKVNMNDLFFSPSEDNTNDLNGEETSDNTNEAEENISLDLQOYLTFFNDPEN 513
DB |
QY 61 ISENLSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGSRILATNSVNE 120
CC |
CC 514 ISENLSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGSRILATNSVNE 573
DB |
QY 121 ALLNPSRVYTFSSDYVKKVNKATEAMFLGWVEQVLYDFTDETSEVSTTDKIADITITII 180
CC |
CC 574 ALLNPSRVYTFSSDYVKKVNKATEAMFLGWVEQVLYDFTDETSEVSTTDKIADITITII 633
DB |
QY 181 PYIGPALNIGNMLYKDDFVGALIFSGA 207
CC |
CC 634 PYIGPALNIGNMLYKDDFVGALIFSGA 660
DB |

```

RESULT 2

```

BXA2_CLOBO STANDARD; PRT; 1295 AA.
ID BXA2_CLOBO
AC Q45894; P77780;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-
GN chain; Botulinum neurotoxin A, heavy-chain].
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN 11]
RP SEQUENCE FROM N.A.
RC STRAIN-KYOTO-F;
RX MEDLINE=9413603; PubMed=8310180;
RA Williams A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with
RT other Clostridial neurotoxins.";
RL Res. Microbiol. 144:547-556(1993).
RN 12]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN-KYOTO-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-gln-1-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (by similarity).
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.

```

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sdb.ch).

```

DR EMBL; X73423; CAA51824.1; -.
DR EMBL; X87974; CAA61234.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro; IPR000395; Bontoxilysin.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR ProDom; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
KM Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1
FT CHAIN 448 1295
FT METAL 222 222
FT ACT_SITE 223 223
FT METAL 226 226
FT DISULFID 429 453
FT DISULFID 1234 1279
FT TRANSMEM 626 646
FT TRANSMEM 655 675
SQ SEQUENCE 1295 AA; 149279 MW; 5DA04A1D98D6372 CRC64;

```

Query Match 85.7%; Score 918; DB 1; Length 1295;

Best Local Similarity 85.4%; Pred. No. 3.3e-66;
Matches 176; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

```

QY 1 IKVNMNDLFFSPSEDNTNDLNGEETSDNTNEAEENISLDLQOYLTFFNDPEN 60
DB |
DB 454 IKVNMNDLFFSPSEDNTNDLNGEETSDNTNEAEENISLDLQOYLTFFNDPEN 513
QY 61 ISENLSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGSRILATNSVNE 120
DB |
DB 514 ISENLSSDIIGOLELMPNIEFPNGKKYELDKYTMFHYLRAQEFHGSRILATNSVNE 573
QY 121 ALLNPSRVYTFSSDYVKKVNKATEAMFLGWVEQVLYDFTDETSEVSTTDKIADITITII 180
DB |
DB 574 ALLNPSRVYTFSSDYVKKVNKATEAMFLGWVEQVLYDFTDETSEVSTTDKIADITITII 633
QY 181 PYIGPALNIGNMLYKDDFVGALIFSG 206
DB |
DB 634 PYIGPALNIGNMLYKDDFVGALIFSG 659

```

RESULT 3

```

BXG_CLOBO STANDARD; PRT; 1296 AA.
ID BXG_CLOBO
AC Q60393;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
GN BONG.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN 11]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for Clostridium botulinum

```

```

RT (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins."
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN PRESYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74162; CAA52275.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn.MTpeptidse.
DR Pfam: PF01742; Peptidase_M27.1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INIT MET 0 BY SIMILARITY.
FT CHAIN 1 441 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT METAL 442 1296 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 435 449 INTERCHAIN (PROBABLE).
SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F665C31 CRC64;
Query Match 35.2%; Score 377; DB 1; Length 1296;
Best local similarity 41.3%; Pred. No. 1.3e-22;
Matches 88; Conservative 34; Mismatches 75; Indels 16; Gaps 3;

```

```

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (Bont/E)
DE (Bontoxilysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43181, and ATCC 43755;
RX MEDLINE=92181428; PubMed=15434481;
RA Poullet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.;
RA Yokosawa N., Tashiki T., Oguma K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.
RC STRAIN=5262;
RA Gimenez J., Foley J., Dasgupta B.R.;
RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
RT partial sequence and comparison.";
RL FASEB J. 2:41750-41750(1988).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN PRESYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62088; CAA43998.1; -.
DR EMBL: X53180; CAA37321.1; -.
DR PIR: JH0256; JH0256.
DR PIR: S16145; S16145.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn.MTpeptidse.
DR Pfam: PF01742; Peptidase_M27.1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW

```

FT INIT_MET 0 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
 FT AC1_SITE 212 212 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 229 229 K -> M (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C2312857 CRC64;

Query Match 35.0%; Score 374.5; DB 1; Length 1250;
 Best Local Similarity 38.9%; Pred. No. 1.9e-22;
 Matches 84; Conservative 51; Mismatches 60; Indels 21; Gaps 7;

OY 1 IKVNMMDLFESPSENFNTND-LINKGEI-----TSQNTIAAEINSLDIQYILFNTD 55
 DB 426 IETNNGELFFVASSENSYNDDNINTFKEIDVTSTNNVE-----NLDQVILNFSE 477
 OY 56 NEPENISIEINLSDDIIGOLELMPNIEFPNG---KVELDRYTFPHYRAQEFHGKSR 111
 DB 478 SAP-GISDEKMLTQIND-AVLPKYD--SNGTSDIEQHVNELNVEFYLDKRPGEEN 533
 OY 112 IALTSTVNEALLNPSRVYTFPSSDYKVKYKATEAPMLGWEQVLVYDFTSEVSTTD 171
 DB 534 VNLTSIDPALLEOPRIYTFPSSSEFINNNKPVQALFVGWIOQVLVDFTTANOKSTVD 593
 OY 172 KIADITIIPIYIGPALNIGNMLYKDFGALIFSG 207
 DB 594 KIADISIVPIYIGLALNIGNEAKGNFKDALELIGA 629

RESULT 5
 BXE_CLOBO STANDARD; PRT; 1250 AA.
 AC 000496:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 DE (Bontolysin E)
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beluga.
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butylicum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92174922; PubMed=1541280;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type-E
 RT neurotoxin, derived by nucleotide sequence analysis of the encoding
 RT gene.";
 RL Eur. J. Biochem. 204:657-667(1992).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=90264400; PubMed=2160960;
 RA Birz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [4]
 RP SEQUENCE OF 1-13
 RX MEDLINE=85197963; PubMed=388113;
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and

RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [5]
 RP SEQUENCE OF 419-426.
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 RT "Botulinum neurotoxin type E fragmented with endoprotease Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217(1990).
 RN [6]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Birz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-1-ILE-
 CC 181 BOND IN SNAP-25.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X62089; CAA43999.1; -;
 DR EMBL; X62683; CAA44558.1; -;
 DR PIR; A60027; A60027.
 DR PIR; B35294; B35294.
 DR PIR; JH0257; JH0257.
 DR PIR; S08575; S08575.
 DR PIR; S18111; S18111.
 DR PIR; S21178; S21178.
 DR HSSP; P10845; 3BPA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontolysin.
 DR InterPro; IPR000130; Zn_Mpeptidase.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOLYSIN.
 DR PRODOM; PD001963; Bontolysin.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0 0
 FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).

```
FT ACT_SITE 212 212 BY SIMILARITY.
FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 411 425 INTERCHAIN (PROBABLE).
FT CONFLICT 176 176 R -> G (IN REF. 2).
FT CONFLICT 197 197 C -> S (IN REF. 2 AND 3).
FT CONFLICT 339 339 R -> A (IN REF. 2).
FT CONFLICT 772 772 I -> L (IN REF. 2).
FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
FT CONFLICT 966 966 R -> A (IN REF. 2).
FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041EB4 CRC64;

Query Match 34.4%; Score 368.5; DB 1; Length 1250;
Best Local Similarity 38.4%; Pred. No. 5,8e-22;
Matches 83; Conservative 51; Mismatches 61; Indels 21; Gaps 7;

OY 1 IKVNNNDLFSPEDNFTND-LNKGEI-----TSDTNEAEENISLIDIQYLFNFND 55
DB 426 IEINNGELFFVASSENSYNDNDNITPKEDDYTSNNNYE-----NDLDQYILNFNSE 477
OY 56 NEPENISINLSDDIGOLELMPNIEFPNG---KKYELDKYTMFHYLRAQEFHGKSR 111
DB 478 SAP-GLSDKRLNTIQND-AIYPKYD--SNQTSIDIOHDVNEINLVEFYLDQKVPGEEN 533
OY 112 IATNVSVMALNPSRVYFFFSDDYKYNKATKAEAMFLGWVQLVYDFDSEVSTTD 171
DB 534 VNTSSIDPALLEQPKRYFFESSEFINNVKPYQALFVSWIQOVLVDFTTEANQSTVD 593
OY 172 KIADITIIIPYIGPALNIGNMLTKDQVGLIFSGA 207
DB 594 KIADISIVPYIGLALNIGNEAKGNFKDALLELGA 629

RESULT 6
BXE CLOBO STANDARD: PRT; 1290 AA.
AC P10844; P10843;
DT 01-JUL-1988 (rel. 11, Created)
DT 01-JUL-1993 (rel. 26, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
DE (Bontoxilysin B).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;
RA Wheelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence.";
RL Appl. Environ. Microbiol. 58:2345-2354(1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN-NCTC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN-NCTC 7273;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
```

```
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin.";
RL Biochimie 70:811-817(1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Sathyanoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E.";
RL Arch. Biochem. Biophys. 238:544-548(1985).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin.";
RL Nature 359:832-835(1992).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
CC SYNAPTOBREVIN-2.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sib.ch).
CC
CC EMBL: M81186; AAA3211.1; -
CC EMBL: Z11934; CAAT7991.1; -
CC EMBL: X70817; CAA50148.1; -
CC
CC PIR: S07128; S07128.
CC PIR: S07135; S07135.
CC PIR: S08562; S08562.
CC PIR: S08573; S08573.
CC PIR: S08574; S08574.
CC PIR: A48940; A48940.
CC HSSP: P10845; 3BTA.
CC MEROPS: M27.002; -.
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_MTPeptidse.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT MET 0 0
FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
```


FT	CHAIN	441	1290	BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
FT	METAL	229	229	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	230	230	BY SIMILARITY.
FT	METAL	233	233	ZINC (CATALYTIC) (BY SIMILARITY).
FT	DISULFID	436	445	INTERCHAIN (PROBABLE).
FT	CONFLICT	29	29	T -> M (IN REF. 4).
FT	CONFLICT	217	217	R -> G (IN REF. 2).
FT	CONFLICT	224	224	A -> S (IN REF. 2).
FT	CONFLICT	463	463	S -> R (IN REF. 4).
SO	SEQUENCE	1290 AA;	150670 MW;	D21745E2C024D43 CRC64;
Query Match		34.2%;	Score 366.5;	DB 1; Length 1290;
Best Local Similarity		39.4%;	Pred. No. 8.8e-22;	
Matches	87;	Conservative	39;	Mismatches 62; Indels 33; Gaps 4;
QY	1	IKVNMWDLFFSPSEDNFTNDLNKGEIITSQDTNIEAAENISLDL1QOYYLFFNFDPEN	60	
Db	446	IDVDNEDELFFIADKNSFDDLSKNERIEYNT-----QSNVIENDF--PIN	488	
QY	61	ISIEMLSDIIGOLEL-----MNIERFPNGKYYELDKYTMHYLAQEEF	106	
Db	489	ELI--LDWDLISKILPSENESLTNDFNVDVPEYVEKOPAIKIFIDENTIFQYLYSQTFP	546	
QY	107	HGKSRIATLNSVNEALLNPSRYTFEPSSDYKVKAKAEAMFLGMEVOLYVDFPENSE	166	
Db	547	LDIRISITLSTSFDDDLNLSNKKYSFSSMDYITANKVYVAGLGLAGMVKQIVNDFEIAEK	606	
QY	167	VSTDKIADIITIIPIYIGPALNIGMLYKDDFVGALLIFSQA	207	
Db	607	SNTMDKIADISLIVPIYIGLALNVGNETPAKGNENMFETAGA	647	
RESULT 7				
BXF_CLOBO				
ID	BXF_CLOBO	STANDARD:	PRT:	1274 AA.
AC	P30996:			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Botulinum neurotoxin type F precursor (BC 3.4.24.69) (BONT/F)			
GN	BONT.			
OS	Clostridium botulinum.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1491;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 23387;			
RX	MEDLINE=93012902; PubMed=1398040;			
RA	East A.K., Richardson P.T., Allaway D., Collins M.D.,			
RA	Roberts T.A., Thompson D.E.;			
RT	"Sequence of the gene encoding type F neurotoxin of Clostridium			
RL	botulinum ";			
RL	FEMS Microbiol. Lett. 75:225-230(1992).			
RN	[2]			
RP	SEQUENCE OF 1-64 FROM N.A.			
RC	STRAIN-Hobbs FT10;			
RX	MEDLINE=94297488; PubMed=7764998;			
RA	East A.K., Collins M.D.;			
RT	"Conserved structure of genes encoding components of botulinum			
RT	neurotoxin complex M and the sequence of the gene coding for the			
RT	neurotoxic component in nonproteolytic Clostridium botulinum type F.";			
RL	Curr. Microbiol. 29:69-77(1994).			
RN	[3]			
RP	SEQUENCE OF 634-1002 FROM N.A.			
RX	MEDLINE=94013372; PubMed=8405842;			
RA	Campbell K., East A.K., Collins M.D.;			
RT	"Gene probes for identification of the botulinum neurotoxin gene and			
RT	specific identification of neurotoxin types B, E, and F.";			
RL	J. Clin. Microbiol. 31:2255-2262(1993).			
RN	[4]			
RP	IDENTIFICATION OF SUBSTRATE.			

```

RA MEDLINE=94230352; PubMed=81755689;
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,
RA Rogues B., Eykhe E.M., Suehof T.C., Jahn R., Niemann H.,
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinum neurotoxins and tetanus toxin".
RL J. Biol. Chem. 269:12764-12772(1994).
CC -1- FUNCTION: BOTULINUM TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT OF THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
CC BOND OF SYNAPTOBREYVINS-1 AND -2.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIANE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M92906; AAA23263.1; -
CC EMBL: S73676; AAC60475.1; -
CC EMBL: X70820; CAA50151.1; -
CC EMBL: X70816; CAA50147.1; -
CC HSSP: P10845; 3BTA.
CC MEROPS: M27.002; -.
CC InterPro: IPR0000395; Bontoxilysin.
CC InterPro: IPR000130; Zn.M1peptide.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; Bontoxilysin.
CC PRODOM: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC-PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolyse; Metalloprotease; Zinc.
CC CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
CC CHAIN 437 1274 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
CC METAL 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACET SITE 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
CC DISULFID 429 445 INTERCHAIN (PROBABLE).
CC SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 29.5%; Score 316; DB 1; Length 1274;
Best Local Similarity 36.2%; Pred. NO. 1e-17;
Matches 79; Conservative 51; Mismatches 62; Indels 26; Gaps 11;

1 IKVNWMDLFFFSSEENFT-NDLNGEELTSDPTNIEAARENSLDLQOYLTFPNNDNEPE 59
Db 1 KVNWMDLFFVAESSSEYENDINPKREDIDPTNNLNNTN-RLD--EVLLDYSNQTIPQ 501
446 IRVNSSELFVAESSSEYENDINPKREDIDPTNNLNNTN-RLD--EVLLDYSNQTIPQ 501
60 NISINLSDDIIGOLELNPNIERFPNG-----KYELDYTFMHYLRADQEPFGSRIALT 115
Db 502 -ISNRTLMT-LVQDSYPRD--SNGSSELEEDYVDENFFYLHQAQVGEGETNISLT 557
OY 116 NSVNEALLNPSRYTFSSSDYVKYVKATTEAAMELGWEOLVYDFTDSEVSTDKIAD 175
Db 558 SSIIDTALLEEK-DIFFSESEFIDTINKPVNALFTIDWISKYIRDFTTETATKSTVDKIAD 616
OY 176 ITIITPYGPAINT-----GNMLAKDDE----GVALI 203
Db 617 ISLVIPYGLANITIEAKNG--FEELFELLGAGILL 652

```

ID	TEXT	CLOTE	STANDARD:	PRT:	1314	AA.
AC	P04958;					
DT	13-AUG-1987 (Rel. 05, Created)					
DT	13-AUG-1987 (Rel. 05, Last sequence update)					
DT	15-JUL-1999 (Rel. 38, Last annotation update)					
DE	Tetanus toxin precursor (EC 3.4.24.68) (Tentoxylysin).					
OC	Clostridium tetani.					
OC	plasmid.					
OC	Bacteria: Firmicutes; Clostridia; Clostridiales; Clostridiaceae;					
OC	Clostridium.					
OX	NCBI_TaxID=1513;					
OX	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=87053814; PubMed=3536478;					
RA	Eisel U., Jarasch W., Goretzki K., Henschen A., Engels J.,					
RA	Weiler U., Hudel M., Habermann E., Niemann H.;					
RT	"Tetanus toxin: primary structure, expression in E. coli, and					
RL	homology with botulinum toxins.";					
RL	EMBO J. 5:2495-2502(1986).					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=CN3911;					
RC	MEDLINE=87040747; PubMed=3774547;					
RA	Fairweather N.F., Lyness V.A.;					
RA	"The complete nucleotide sequence of tetanus toxin.";					
RL	Nucleic Acids Res. 14:7809-7812(1986).					
RP	SEQUENCE OF 742-1314 FROM N.A.					
RX	MEDLINE=86085672; PubMed=3510187;					
RA	Fairweather N.F., Lyness V.A., Pickard D.J., Allen G., Thomson R.O.;					
RT	"Cloning, nucleotide sequencing, and expression of tetanus toxin					
RL	fragment C in Escherichia coli.";					
RL	J. Bacteriol. 165:21-27(1986).					
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.					
RX	MEDLINE=90201034; PubMed=2108021;					
RA	Kriegstein K., Henschen A., Weiler U., Habermann E.;					
RT	"Arrangement of disulfide bridges and positions of sulfhydryl groups					
RL	in tetanus toxin.";					
RL	Eur. J. Biochem. 188:39-45(1990).					
RP	PARTIAL SEQUENCE.					
RX	MEDLINE=92037649; PubMed=1935979;					
RA	Kriegstein K.G., Henschen A.H., Weiler U., Habermann E.;					
RT	"Limited proteolysis of tetanus toxin. Relation to activity and					
RL	identification of cleavage sites.";					
RL	Eur. J. Biochem. 202:41-51(1991).					
RP	IDENTIFICATION AS ZINC-PROTEASE.					
RX	MEDLINE=93010948; PubMed=1396588;					
RA	Schavo G., Poulin B., Rossetto O., Benfenati F., Tauc L.,					
RA	Montecucco C.;					
RT	"Tetanus toxin is a zinc protein and its inhibition of					
RT	neurotransmitter release and protease activity depend on zinc.";					
RL	EMBO J. 11:3577-3583(1992).					
RP	IDENTIFICATION OF SUBSTRATE.					
RX	MEDLINE=93063293; PubMed=1331807;					
RA	Schavo G., Benfenati F., Poulin B., Rossetto O., de Lauro P.P.,					
RA	Dasgupta B.R., Montecucco C.;					
RT	"Tetanus and botulinum-B neurotoxins block neurotransmitter release					
RT	by proteolytic cleavage of synaptobrevin.";					
RL	Nature 359:852-855(1992).					
RP	X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 874-1314.					
RX	MEDLINE=97475217; PubMed=9334741;					
RA	Umland T.C., Wingert L.M., Swaminathan S., Furey W.F., Schmidt J.J.,					
RA	Sax M.;					
RT	"Structure of the receptor binding fragment HC of tetanus					

RESULT 9
 BXL_CLOBO STANDARD; PRT; 1276 AA.
 AC P19321;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BONT/D)
 DE (bontoxilysin D).
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BVD/-3;
 RX MEDLINE=91016853; PubMed=2216736;
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,
 RA Kozaki S., Kriegstein K., Henschen A., Gili D.M., Niemann H.,
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum
 RT neurotoxin type D.";
 RL Nucleic Acids Res. 18:5556-5556(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CB16;
 RX MEDLINE=93042276; PubMed=1420572;
 RA Sunagawa H., Ohnaka T., Watanabe T., Inoue K.;
 RT "The complete amino acid sequence of the Clostridium botulinum type D
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT phase d-16 phi genome.";
 RL J. Vet. Med. Sci. 54:905-913(1992).
 RN [3]
 RP PARTIAL SEQUENCE.
 RC STRAIN-D-SA, and D-1873;
 RX MEDLINE=89339741; PubMed=2668193;
 RA Morishiki K., Syuto B., Kubo S., Oguma K.;
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D
 RT strains.";
 RL Infect. Immun. 57:2886-2891(1989).
 RN [4]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94230352; PubMed=8175689;
 RA Yamasaki S., Baumeister A., Binz T., Biasi J., Link E., Cornille F.,
 RA Regue B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
 RT F botulinum neurotoxins and tetanus toxin.";
 RL J. Biol. Chem. 269:12764-12772(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-1-LEU-61 BOND OF
 CC SYNAPTOSOMES-1 AND -2.
 CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 CC BACTERIOPHAGE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL; X54254; CAA38175.1; -;
 DR EMBL; S49407; AAB24244.1; -;
 DR PIR; S11455; S11455.
 DR HSP; P10845; 3BTA.
 DR MEROPS; M27.002; -;
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_MTPeptide.
 DR Pfam; PF01742; Peptidase_M27_1.
 DR PRINTS; PR00760; BONTOTOXILYSIN.
 DR PRODOM; PD001963; Bontoxilysin; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolyase; Zinc.
 FT CHAIN 1 442
 FT METAL 443 1276
 FT METAL 229 229
 FT ACT_SITE 230 233
 FT METAL 233 233
 FT DISULFID 437 450
 FT VARIANT 15 16
 FT VARIANT 17 18
 FT VARIANT 452 452
 FT VARIANT 457 457
 FT VARIANT 457 457
 FT VARIANT 462 462
 FT VARIANT 489 489
 FT VARIANT 644 644
 FT VARIANT 1122 1122
 SQ SEQUENCE 1276 AA; 146871 MW; C1EC50F4C8233E2 CRC64;
 Query Match 27.1%; Score 290.5; DB 1; Length 1276;
 Best Local Similarity 32.6%; Pred. No. 1,1e-15;
 Matches 71; Conservative 39; Mismatches 79; Indels 29; Gaps 5;
 QY 1 IKVNNMDLFFSPSEBNTNDLNGKEITSPTNIEAENISLDLIDQYVLTENDEPEN 60
 DB 451 IKVNNMLPYVADKDSISQEIFENKILTDETNQVNSDKFSLD--ESILDGVPIINPEI 507
 QY 61 ISIEINSSDILIGLELMPNIERP-----NGKYEYEDKYMFLRAOEREHG 108
 DB 508 VD-----PLTPNVMPELNIPEEIVFYDDITY-VDYNSYLLYSQKISNN 554
 QY 109 KSRIALTNSVNEALLNPSRYTFESSDYKVKKATEAAMFLGWEDLVYDFTDESEVS 168
 DB 555 VENTILTVTSVEALGYSNKTITFLPS-LAEKVKGVAGGLFWMANGVEDFTTNIMKKD 613
 QY 169 TTDKIADITITIPYIGPALNIGMKLYKDFVGLITFSG 206
 DB 614 TLDKISDVSVIIPYIGPALNIGNSALGNQAFATAG 651
 RESULT 10
 BXL_CLOBO STANDARD; PRT; 1290 AA.
 AC P18640;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type C1 precursor (EC 3.4.24.69) (BONT/C1)
 DE (bontoxilysin C1).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90370487; PubMed=2204031;
 RA Hauser D., Eklund M.W., Kurazono H., Binz T., Niemann H., Gili D.M.,

RA Boquet P., Popoff M.R.:
 "Nucleotide sequence of Clostridium botulinum C1 neurotoxin.";
 Nucleic Acids Res. 18:4924-4924(1990).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-TYPE C Stockholm / C-ST:
 RX MEDLINE=91024998; PubMed=2222445;
 RA Kimura K., Fujii N., Tsuruki K., Murakami T., Indoh T.,
 RA Yokosawa N., Takeshi K., Syuto B., Oguma K.:
 "The complete nucleotide sequence of the gene coding for botulinum
 type C1 toxin in the C-ST phage genome.";
 RT Biochem. Biophys. Res. Commun. 171:1304-1311(1990).
 [3]
 RN SEQUENCE OF 2-25
 RC STRAIN-TYPE C Stockholm / C-ST:
 RX MEDLINE=88153072; PubMed=2450068;
 RA Tsuruki K., Yokosawa N., Syuto B., Ohishi I., Fujii N., Kimura K.,
 RA Oguma K.:
 "Establishment of a monoclonal antibody recognizing an antigenic site
 common to Clostridium botulinum type B, C1, D, and E toxins and
 tetanus toxin.";
 RT Infect. Immun. 56:898-902(1988).
 [4]
 RN IDENTIFICATION OF SUBSTRATE.
 RA MEDLINE=94038966; PubMed=7901002;
 RA Blas J., Chapman E.R., Yamasaki S., Binz T., Niemann H., Jahn R.:
 "Botulinum neurotoxin C1 blocks neurotransmitter release by means of
 cleaving HPC-1/syntaxin.";
 RT EMBO J. 12:4821-4828(1993).
 CC -I- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 ENDOPEPTIDASE THAT CLEAVES SYNTAXIN.
 CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 detected action on small molecule substrates.
 CC -I- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -I- MISCELLANEOUS: BOTULINUM TYPE C1 NEUROTOXIN IS SYNTHESIZED BY C
 STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE
 BACTERIOPHAGE.
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X66433; CAA47060.1; -
 DR EMBL; X72793; CAA51313.1; -
 DR EMBL; X53751; CAA37780.1; -
 DR EMBL; D90210; BAA14235.1; -
 DR EMBL; X62389; CAA44263.1; -
 DR PIR; S11291; S11291.
 DR PIR; A35396; A35396.
 DR PIR; A43503; A43503.
 DR HSSP; P10845; 3BTA.
 DR MEROPS; M27.002; -
 DR InterPro; IPR000395; Bontoxilysin.
 DR InterPro; IPR000130; Zn_MTPeptidase.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTOXILYSIN.
 DR ProDom; PD001963; Bontoxilysin; 1.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT_MET 0 0
 FT CHAIN 1 448 BOTULINUM NEUROTOXIN C1, LIGHT-CHAIN.
 FT CHAIN 449 1290 BOTULINUM NEUROTOXIN C1, HEAVY-CHAIN.
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 229 229 BY SIMILARITY.
 FT METAL 232 232 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFD 436 452 INTERCHAIN (PROBABLE).
 FT CONFLICT 84 84 P -> T (IN REF. 2).
 SQ SEQUENCE 1290 AA; 148734 MW; 71FBE379F97129E8 CRC64;
 Query Match 24.6%; Score 263; DB 1; Length 1290;
 Best Local Similarity 31.2%; Pred. No. 1.9e-13;
 Matches 64; Conservative 39; Mismatches 94; Indels 8; Gaps 4;
 QY 3 VANNMLFSPSDNFTNDLNKGEITSPNIFAENISLDIQOYVTFNPNDEPNIS 62
 DB 457 VAKNTDLPIFGDISDVKTDFLRKADINEETVYIDNVSVD---QVILS---KNISEGQ 510
 QY 63 IENLSDDIIQLELMPNIEP-FPNGKYYELDKYTFEHLRAQEFHGRSRLATNSVNEA 121
 DB 511 IDLVPSIDSESEILPGENOVFYDNRTQNVDLNYSYLSQKLSDNVEDFTFTRISEA 570
 QY 122 LNPESRYVTFESDVKYKNTKTEAMFLGWQELVYFTDTSVSTDKTADITITIP 181
 DB 571 IDNSAKVYTYFPPT-LANNVAGVGGGLFLMANDVEDFTNIIKRDTLKDIDVSAIIP 629
 QY 182 YIGPALNIGNMLYKDDFVGAIFSG 206
 DB 630 YIGPALNISNSVRRONFTFAFVATG 654
 RESULT 11
 Y402_BUCAI STANDARD; PRT; 246 AA.
 ID Y402_BUCAI
 AC P57482;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BU402.
 GN BU402.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Tokyo 1998;
 RA MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.:
 "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. Aps.";
 RT Nature 407:81-86(2000).
 RL Nature 407:81-86(2000).
 CC -I- SIMILARITY: BELONGS TO THE UPF0169 (COML) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP001119; BAB13105.1; -
 DR InterPro; IPR005156; UPF0169.
 DR Pfam; PF03696; UPF0169; 1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 246 AA; 29937 MW; 1289ED7ACD94088F CRC64;
 Query Match 10.0%; Score 107; DB 1; Length 246;
 Best Local Similarity 21.8%; Pred. No. 0.088;
 Matches 44; Conservative 40; Mismatches 74; Indels 44; Gaps 8;

[illegible]

```

01 143 ATEAMFLGAWBOLVYDPTDENFSEVSTDPKADIILIIIPYIGPAINIGMLKKDPVGL 202
02 553 DTDKRYTL-WLKEVKNKNSFDINLTQEIIDSMGCIINEVYLMFKALNLTNT--SNSFVEEY 609
03 203 IFSGA 207
04 610 QDSGA 614
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017

```

DB 156 TPLAPYKNCFDRAEL-DMNIEIINKLYKAYLEDFYVTEIEPEPAKECQTLIGF 214
 QY 53 NFDNEPENISIEML-SSDIIGOL--ELMPNIEEPNGKYEKDYTMFHYLRAQEFHKG 109
 DB 215 EADRRSINITALNLSQSSDIDPDLSKDLPLN-----GKLVPL--ATPFLAQAQDFGVR 266
 QY 110 SRIA 113
 DB 267 AALA 270

RESULT 14
 MT18_YEAST
 ID MT18_YEAST STANDARD; PRT; 1032 AA.
 AC P40469; P89106;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE DNA repair/transcription protein MET18/MMS19.
 GN MET18 OR MMS19 OR YIL128W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 NC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RX MEDLINE=97098656; PubMed=8943333;
 RA Lauder S., Bankmann M., Guzder S.N., Sung P., Prakash L.,
 RA Prakash S.;
 RT "Dual requirement for the yeast MMS19 gene in DNA repair and RNA
 RT polymerase II transcription.";
 RL Mol. Cell. Biol. 16:6783-6793(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-161 FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Churcher C., Barrall B.G., Rajandream M.A.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 162-1032 FROM N.A.
 RC STRAIN=S288c / AB972;
 RA Hamlyn N., Churcher C., Barrall B.G., Rajandream M.A.;
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN NUCLEOTIDE EXCISION REPAIR (NER) AND RNA
 CC POLYMERASE II (POL II) TRANSCRIPTION. IT PROBABLY DOES NOT
 CC PARTICIPATE DIRECTLY IN NER AND POL II TRANSCRIPTION BUT EXERTS
 CC ITS BIOLOGICAL EFFECTS BY INFLUENCING THE ACTIVITY OF TFIIH AND
 CC POSSIBLY OTHER DNA REPAIR AND TRANSCRIPTION FACTORS AS AN UPSTREAM
 CC REGULATORY ELEMENT. INVOLVED IN SPORULATION AND POSSIBLY IN THE
 CC RAD52 RECOMBINATIONAL REPAIR PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DOMAIN: HAS 15 TANDEN LEUCINE RICH REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U70559; AAB38865.1; -
 DR EMBL: Z38059; CAA86150.1; -
 DR EMBL: Z46833; CAA86864.1; -

DR TRANSFAC; T03493; -
 DR SGD; S0001390; MET18.
 KW DNA repair; Nuclear protein; Repeat.
 FT CONFLICT 230 230 P -> Q (IN REF. 1).
 FT CONFLICT 329 329 D -> G (IN REF. 1).
 FT CONFLICT 335 335 V -> M (IN REF. 1).
 FT CONFLICT 361 361 V -> I (IN REF. 1).
 SQ SEQUENCE 1032 AA; 117882 MW; D05DC488E098814 CRC64;

Query Match 8.9%; Score 95.5; DB 1; Length 1032;
 Best Local Similarity 22.1%; Pred. No. 4.5;
 Matches 46; Conservative 33; Mismatches 76; Indels 53; Gaps 10;

QY 3 VNMMDFEESP-----SEDNFTNDLNK--GEITSDTNIEAEENISLDLIQY 48
 DB 297 LEWMTLLMNLKFEIMQNSGENETLLNPNKQSDVDGYNTYDACLKIINMALQLY 356
 QY 49 YLTFNPNENPENISIEMLSSDIIGOLELMNIEERFPGKYEKDYTMFHYLRAQEFHKG 108
 DB 357 ---NFD---KVSFEKFFTHVD--ELKPNF-----KYEDLKQTCQILSA---IG 395
 QY 109 KSRIALNVSNEA-----LLNPSRV-----YFPSSDYKKVKKATEAMFLGWE 154
 DB 396 SGNVEIRNKYISSTFPLFLINTSEVAKLLIMNFSEFVSYIDLRGRTSKESLGTPVPN 455
 QY 155 QLVYDFPDE---TSEVSTPDKIADITI 178
 DB 456 NKMAEYKDEIMLSMALTRSSKAKEYTI 483

RESULT 15
 Y198_RICPR
 ID Y198_RICPR STANDARD; PRT; 355 AA.
 AC 092DM7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein RP198 precursor.
 GN RP198.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.B., Zomorodipour A., Andersson J.O.,
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Neslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria.";
 RL Nature 396:133-140(1998).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ235270; CAA14663.1; -
 CC Hypothetical protein; Signal; Complete proteome.
 KW SIGNAL 1 21
 FT CHAIN 22 355 HYPOTHETICAL PROTEIN RP198.
 SQ SEQUENCE 355 AA; 40780 MW; A08B7655BCDBE0 CRC64;

Query Match 8.7%; Score 93.5; DB 1; Length 355;
 Best Local Similarity 26.4%; Pred. No. 1.7;
 Matches 57; Conservative 32; Mismatches 76; Indels 51; Gaps 14;

QY 1 IKVNMNDL-FFSPSEDNF-TNDLNKGEITSDTNIEAEENISLDLIQY YLTFNPNDEP 58

```

Db 146 IKLNKTLISIFKONQEIFKINDL---AFILTKHNL-ASQENISLFLMHYY-----SEK 195
QY 59 ENISIEMLSDIIGOLELMP-----NIRF-----PNGKKYELDKYTMFHYLRAQ 103
Db 196 DILNFKNANLDMATSFPAKNGKDAILENLIERFIETCDNESKVNINGTLOFF---AN 252
QY 104 EFHGRSRIALT--NSVNEALLNPSRYTFESSDYKKV-----NKATEAAMFLGWVEOL 156
Db 253 KLPKGIILSFELNYSIVDKIL-PNSI--LFSKKTITIKIYAKAMNKTSD-----EOL 301
QY 157 VYDFDETSEVSTYDKIADITIIIPYIGPALNIGNM 192
Db 302 ---NFDKNDTNSVYNNIKNAKFDIAFSDKGINIGSM 334

```

Search completed: March 13, 2003, 11:40:36
 Job time : 13.5447 secs

THIS PAGE BLANK (USPTO)